



PF 18-JUN-1997; 97WO-SF01091.  
 XX  
 PR 20-JUN-1996; 96SE-0002496.  
 XX  
 PA (FLOC/) FLOCK J.  
 PA (FRK/) FRYKBERG L.  
 PA (GUS/) GUS B.  
 PA (LIND/) LINDBERG M.  
 PA (NILS/) NILSSON M.  
 PI  
 XI Flock J, Frykberg L, Gus B, Lindberg M, Nilsson M;  
 XX MPI: 1998-063079/06.  
 XX N-PSDB: AAV04279.  
 XX  
 PR Fibrinogen-binding protein from coagulase-negative Staphylococcus -  
 PT used for prevention, treatment and diagnosis of Staphylococcus  
 PT Infection  
 XX  
 PS Example 3; Fig 6; 45pp; English.

The protein comprises the fibrinogen binding protein (FIG) of coagulase-negative staphylococcus epidermidis HB. Its amino acid sequence was deduced from the isolated fig gene (see AAV04279). The closest known analogue of FIG is the clumping factor of S. aureus which also binds fibrinogen and promotes bacterial aggregation in serum. Recombinant FIG polypeptides can be expressed in host cells. They are used as immunogens, particularly in vaccines (which may be expressed in vivo) to protect humans and animals against coagulase-negative staphylococcus infection. Antibodies raised against FIG can be used for passive immunisation. They block the adherence of bacteria and for diagnosis.

Sequence 1092 AA;

Query Match 100.0%; Score 5646; DB 19; Length 1092;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-266;  
 Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MINKNNLLTKKKPKANKSNKAIKRFYGTASTIYGATLFLGLCHNEAKAEENSVQDVK 60  
 DB 1 MINKNNLLTKKKPKANKSNKAIKRFYGTASTIYGATLFLGLCHNEAKAEENSVQDVK 60  
 QY 61 DSNTDELSD 120  
 DB 61 DSNTDELSD 120  
 QY 121 TTNDENEAFLQKTPQDNTHLTEBEVKESSESSNSIDTAQOPSHTTINREESVQTS 180  
 DB 121 TTNDENEAFLQKTPQDNTHLTEBEVKESSESSNSIDTAQOPSHTTINREESVQTS 180  
 QY 181 DNEVESHVDFPANSKIKSNSTSGKEENTIDQPNKVKEDSTTSQPSGYTINDEKTSNDE 240  
 DB 181 DNEVESHVDFPANSKIKSNSTSGKEENTIDQPNKVKEDSTTSQPSGYTINDEKTSNDE 240  
 QY 241 LLNLPINEYENKARPLSTSAQPSIKRYVNOALAEQGSNVHLIKVDOSTITGEYDSE 300  
 DB 241 LLNLPINEYENKARPLSTSAQPSIKRYVNOALAEQGSNVHLIKVDOSTITGEYDSE 300  
 QY 301 GVKAHDAENLIYDTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIDNGEITA 360  
 DB 301 GVKAHDAENLIYDTFEVDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIDNGEITA 360  
 QY 361 TGTYNKNAKQITTYTVDYDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTALSVNK 420  
 DB 361 TGTYNKNAKQITTYTVDYDKYENIKAHKLKTSYIDKSKVPNNNTKLDVEYKTALSVNK 420  
 QY 421 TITVEYORPENRANTLQSMFTNIDTKNHTVEQTIYINPLRSARETNVNTISGNDSEGT 480  
 DB 421 TITVEYORPENRANTLQSMFTNIDTKNHTVEQTIYINPLRSARETNVNTISGNDSEGT 480  
 QY 481 IIDSTIIKVKVGNQNLPSNRITDYSEEDVTNDDYAQLGNNDVNINFGNIDSPFI 540  
 DB 481 IIDSTIIKVKVGNQNLPSNRITDYSEEDVTNDDYAQLGNNDVNINFGNIDSPFI 540

DB 481 IIDSTIIKVKVGNQNLPSNRITDYSEEDVTNDDYAQLGNNDVNINFGNIDSPFI 540  
 QY 541 IKVISKYPNKKDYTTIOQVTYMTQTTINEYTGFEFTASDYNTIARSTSGGQGGDLPEPK 600  
 DB 541 IKVISKYPNKKDYTTIOQVTYMTQTTINEYTGFEFTASDYNTIARSTSGGQGGDLPEPK 600  
 QY 601 TYKIDYVMEVDYDKGIONTNDNEKPLSNVLTLYTPDGTSAKSVYTTDEGKQFPGKLG 660  
 DB 601 TYKIDYVMEVDYDKGIONTNDNEKPLSNVLTLYTPDGTSAKSVYTTDEGKQFPGKLG 660  
 QY 661 LTYKITFEETPEGYPTTLKHSFTNPLADSEGNVWYTINGODMTIDSGFYQTPKISLGNY 720  
 DB 661 LTYKITFEETPEGYPTTLKHSFTNPLADSEGNVWYTINGODMTIDSGFYQTPKISLGNY 720  
 QY 721 VWYDTNKGIOGDEKGISGVKVTLKDENGNIISTTTDENGKYQFDNLNSGNYIVHEDK 780  
 DB 721 VWYDTNKGIOGDEKGISGVKVTLKDENGNIISTTTDENGKYQFDNLNSGNYIVHEDK 780  
 QY 781 PSGMTQTTTSDGDDDEQADGEEVAVTTTDHDFSIDNGYDDESDSDSDSDSDSDSDSD 840  
 DB 781 PSGMTQTTTSDGDDDEQADGEEVAVTTTDHDFSIDNGYDDESDSDSDSDSDSDSDSD 840  
 QY 841 SD 900  
 DB 841 SD 900  
 QY 901 SD 960  
 DB 901 SD 960  
 QY 961 SD 1020  
 DB 961 SD 1020  
 QY 1021 SD 1080  
 DB 1021 SD 1080  
 QY 1081 LGRKRRNRKNK 1092  
 DB 1081 LGRKRRNRKNK 1092

RESULT 2  
 ABP40469  
 ID ABP40469 standard; Protein: 930 AA.  
 XX  
 AC ABP40469;  
 XX  
 DT 24-JUL-2002 (first entry)  
 DE  
 XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.  
 DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 KW antibacterial; gene therapy.  
 OS Staphylococcus epidermidis.  
 XX  
 PN US6380370-B1.  
 PD 30-APR-2002.  
 XX  
 PF 13-AUG-1998; 98US-0134001.  
 XX  
 PR 14-AUG-1997; 97US-055779P.  
 PR 08-NOV-1997; 97US-064964P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Doucetle-Stamm LA, Bush D;  
 XX  
 DR MPI: 2002-381255/41.  
 DR N-PSDB: ABN93014.

XX	Novel isolated nucleic acid encoding a Staphylococcus epidermidis	
PT	polypeptide, useful for diagnosing and treating bacterial infections -	
XX	Disclosure; SEQ ID 5314; 267pp; English.	
CC	ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading	
CC	frame (ORF) nucleic acid sequences which encode the amino acid sequences	
CC	given in ABP95124 to ABP97960. The S. epidermidis sequences have	
CC	antibacterial activity and can be used in gene therapy. The sequences	
CC	can also be used in the diagnosis and treatment of bacterial infections,	
CC	particularly S. epidermidis infections. The sequences can be used to	
CC	screen for compounds able to interfere with the S. epidermidis life	
CC	cycle or inhibit S. epidermidis infection.	
CC	N.B. The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from the	
CC	USPTO web site.	
XX	Sequence 930 AA;	
SO	Query Match 79.5%; Score 4490; DB 23; Length 930;	
	Best Local Similarity 81.2%; Pred. No. 1.9e-210;	
	Matches 884; Conservative 19; Mismatches 26; Indels 160; Gaps 3	
QY	4 KNNLLTKKKRPLANKSNKYAIRKFTVGAASYIGATLLFGLHNEKAKEANSVODYKDSN 63	
DB	2 KNNLLTKKKRPLANKSNKYAIRKFTVGAASYIGATLLFGLHNEKAKEANSVODYKDSN 61	
QY	64 TDDLESDSNDOSDDEKNDVNNNOSINTDDNNQIIRKEETNNYDGIKRSRDRRESTN 123	
DB	62 MDDLESDSNDOSDNEKNDVNNNOSINTDDNNQIIRKEETNNYDGIKRSRDRRESTN 120	
QY	124 VDENEATFLQTPDONTNLTLEEVKESSVESSESSNIDTAQOPSHTTIRRESVQTSNV 183	
DB	121 VDENEATFLQTPDONTNLTLEEVKEPESSVESSESSMDTAQOPSHTTINSEASIQTSNDE 180	
QY	184 EDSHSDSANSIKRSNTEESCKEENTIEOPNKVKEDSTSPSGYTNIDEKISNODELLN 243	
DB	181 ENSRSDSANSIKRSNTEESCKEENTIEOPNKVKEDSTSPSGYTNIDEKISNODELLN 240	
QY	244 LPINEYENKARPLSTTSAPSIKRYTVNOLAEOGSNVNHLIKVPDOSTIEGVDSEGYI 303	
DB	241 LPINEYENKARPLSTTSAPSSKRYTVNOLAEOGSNVNHLIKVPDOSTIEGVDSDGIT 300	
QY	304 KAHDAENLIYDVFEEVDKVKSGDTMTVDIDKNVPSDLTDSFTPIKIKDNGSEIATGT 363	
DB	301 KAHDAENLIYDVFEEVDKVKSGDTMTVDIDKNVPSDLTDSFALPIKIKDNGSEIATGT 360	
QY	364 YDNKKKQITTTTVDYVDKYEENKAHLKLTSYIDKSKVPNNNRKLDVEYETALLSVKTTI 423	
DB	361 YDNKKKQITTTTVDYVDKYEENKAHLKLTSYIDKSKVPNNNRKLDVEYETALLSVKTTI 420	
QY	424 VEYQRPNEKRTANLOSMPFNIDTKNHTVEQTYINLRLYSAKETNNVNIISGNGDEGSTIID 483	
DB	421 VEYQRPNEKRTANLOSMPFNIDTKNHTVEQTYINLRLYSAKETNNVNIISGNGDEGSTIID 480	
QY	484 DSTIIRKVVKGDNQNLPSNRIYDYSEYEDVYNDYDQAOLGNNDVNIINFGNIDSPYIIKY 543	
DB	481 DSTIIRKVVKGDNQNLPSNRIYDYSEYEDVYNDYDQAOLGNNDVNIINFGNIDSPYIIKY 540	
QY	544 ISKYPNPKDDYTTTIOQTYVMQTTINETYGEFPRASVDNTIAPSTSSGGQOQZDLPERKTYK 603	
DB	541 ISKYPNPKDDYTTTIOQTYVMQTTINETYGEFPRASVDNTIAPSTSSGGQOQZDLPERKTYK 600	
QY	604 IGDYWMEDVDKGIQONTNNEKPLSNVLTLYTPDSTSKSVPRDEGCKQOZDLKNGLTYY 663	
DB	601 IGDYWMEDVDKGIQONTNNEKPLSNVLTLYTPDSTSKSVPRDEGCKQOZDLKNGLTYY 660	
QY	664 KITFEPEGEYPTLIKHSQTNPALDSBGSNVWYINGODDWTIDSGFYQTPKYSILGNVWY 723	
DB	661 KITFEPEGEYPTLIKHSQTNPALDSBGSNVWYINGODDWTIDSGFYQTPKYSILGNVWY 720	
QY	724 DTNKGIOGDDEKGISGVAVYTLKDENGNIISTTTDDENGKYQPDNLNSGNVIYHFDKPSG 763	

Db	721	DTNKDGIQGDDEKGISGVKVTLDENGNISITTTDENGKTYQFPDNLNSGNYIVHFDKPSG	780
QY	784	MTQTTTSSGDDDEQADAGEVHYITTDHDFSIIDNGYIDDESDDSDSDSDSDSDSDS	843
Db	781	MTQTTTSSGDDDEQADAGEVHYITTDHDFSIIDNGYIDDESDDSDSDSDSDSDSDS	833
QY	844	DSOSDS	903
Db	834	-----	833
QY	904	DSOSDS	963
Db	834	-----	833
QY	964	DSOSDS	102
Db	834	DSOSDS	860
QY	1024	GSOSDS	108
Db	881	-----SSDKNTKDKLPDYGANEDHDSKGTILGALFAGLGALLGK	921
QY	1084	RRKNRKNKN 1092	
Db	922	RRKNRKNKN 930	
RESULT 3			
ID	AAV83171	standard; Protein; 991 AA.	
AC	AAV83171:		
DT	24-JUL-2000	(first entry)	
DE	Cell wall protein SdrG.		
KW	SdrF, SdrG, SdrH; coagulase negative; staphylococcus; scepticemia;		
KW	osteomyelitis; endocarditis; immune response; vaccine; graft;		
KW	stentyl; intravenous catheter; heart valve; cardiac.		
OS	Staphylococcus sp.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 14	/note- "Position encoded by TAG stop codon"	
FT	Misc-difference 33	/note- "Position encoded by TAG stop codon"	
FT	Misc-difference 964	/note- "Position encoded by TGA stop codon"	
FT	Misc-difference 980	/note- "Position encoded by TAA stop codon"	
FT	Misc-difference 989	/note- "Position encoded by TAG stop codon"	
FT	Misc-difference 989	/note- "Position encoded by TAA stop codon"	
XX	WO200012689-A1.		
XX	09-MAR-2000.		
XX	31-AUG-1999;	99WO-US19728.	
XX	31-AUG-1998;	98US-0098443.	
XX	25-JAN-1999;	99US-0117119.	
XX	(QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.		
XX	(TEXA ) UNITV TEXAS A & M SYSTEM.		
XX	Foster TJ, Hook M, Davis S, Hartford O, McCrea K, NI Eidihi D;		
XX	WPI; 2000-256637/22.		
XX	N-PSDB; AAZ93534.		

0v 724 PBNKDCIOGDNFKCISCYKWTIKNENCNTCTETETEDENVCVQEDNTNCGNVITIHEDKDCG 78

XX

XX Patl Jm, Foster TJ, Hook M;  
PI



DR MPI: 2000-237781/20.  
 DR N-PSDB; AA2511202.  
 XX Composition used for generating immune response or for inhibiting  
 PT microbial colonization in an animal comprises antibodies that bind  
 PT collagen binding protein, fibrinogen binding protein and, optionally,  
 PT fibronectin binding protein.  
 XX  
 XX Claim 8; Fig 4; 115pp; English.  
 XX  
 CC The patent discloses multicomponent vaccines containing selected  
 CC combinations of bacterial binding proteins termed MSCRAMM (microbial  
 CC surface components recognising adhesive matrix molecules) or their  
 CC antibodies. A vaccine composition is provided that includes collagen  
 CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein  
 CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB),  
 CC and optionally a fibronectin binding protein e.g. FBP-A.  
 CC The vaccines are useful for imparting protection against a broad  
 CC spectrum of Staphylococcal strains and for inhibiting microbial  
 CC colonisation, especially of Staphylococcus aureus, in an animal.  
 CC The combinations can also be used to select donor blood pools for the  
 CC preparation of purified blood products for passive immunisation.  
 CC The present sequence is a serine-aspartate repeat region  
 CC protein, SdrC from Staphylococcus epidermidis. The Sdr protein is  
 CC useful in vaccine preparation in combination with specific  
 CC bacterial binding proteins. These vaccines can be used to treat a broad  
 CC spectrum of bacterial infections, including those arising from both  
 CC coagulase-positive and coagulase-negative bacteria.  
 CC  
 XX  
 XX Sequence 991 AA:  
 Query Match 79.4%; Score 4485; DB 21; Length 991;  
 Best Local Similarity 81.1%; Pred. No. 3.6e-210;  
 Matches 883; Conservative 19; Mismatches 27; Indels 160; Gaps 3;

Db 574 ISKYDPKDDYTTIQGTVMQTINNEYTGEPRTASNDNTAFSTSSGOGGDLPEPKTYK 633  
 QY 604 IGDYWEDEVKDGIGQNTNDNEKPLSNVLTLYPDGTSKSVREDEKQYFDGLKNGLT 663  
 Db 634 IGDYWEDEVKDGIGQNTNDNEKPLSNVLTLYPDGTSKSVREDEKQYFDGLKNGLT 693  
 QY 664 KITFETPEGYTPPLKHSCTNPALDSEGNYSWVTINGDDMTIDSGFYQTRKYSIGNTVMY 723  
 Db 694 KITFETPEGYTPPLKHSCTNPALDSEGNYSWVTINGDDMTIDSGFYQTRKYSIGNTVMY 753  
 QY 724 DTNKGIGQDDEKIGSVKYLKDENGNITSTTIDENKGYQPNLNSGNVYHFDKPSG 783  
 Db 754 DTNKGIGQDDEKIGSVKYLKDENGNITSTTIDENKGYQPNLNSGNVYHFDKPSG 813  
 QY 784 MTQTTDSDGDDDEADAGEEYHVTITPDHDFSIDNGYDESDSDSDSDSDSDSDSDSDSD 843  
 Db 814 MTQTTDSDGDDDEADAGEEYHVTITPDHDFSIDNGYDESDSDSDSDSDSDSDSDSDSD 866  
 QY 844 DSD 903  
 Db 867 ----- 866  
 QY 904 DSD 963  
 Db 867 ----- 866  
 QY 964 DSD 1023  
 Db 867 DSD 913  
 QY 1024 GSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOSDSOS 1083  
 Db 914 -----NSSDKTKDKLPDTGANEDHDSKGTLLALFAGLALLGK 954  
 QY 1084 RRRNRKNKN 1092  
 Db 955 RRRNRKNKN 963

RESULT 5  
 AAY08643  
 ID AAY08643 standard; Protein: 1166 AA.  
 XX  
 AC AAY08643;  
 XX  
 DT 09-AUG-1999 (first entry)  
 XX  
 DE S. aureus SdrE protein.  
 XX  
 KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;  
 KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;  
 KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;  
 KW extracellular matrix; vascular graft; vascular stent; vaccine;  
 KW intravenous catheter; artificial heart valve; cardiac assist device;  
 KW antibacterial.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 PN W09927109-A2.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PE 25-NOV-1998; 98NO-US25246.  
 XX  
 PR 31-AUG-1998; 98US-0098427.  
 PR 26-NOV-1997; 97US-0066815.  
 XX  
 PA (EIDH/) EIDHIN D N.  
 PA (FORE-) FORPAS T/A BIORESEARCH IRELAND.  
 PA (FOST/) FOSTER T J.  
 PA (HOOK/) HOOK M A O.  
 PA (INH-) INHIBITEX INC.  
 PA (JOSE/) JOSEFSSON E.

D	b		396	KTPV-NETSLNLFATAGKRTSQNVYVDQPVPHVHDSNIQSIFTKLDBDKTIEQQIVY	454
O	y	458	NPLRYSAKETENVISGNG-----DGGSTIIDSTIIKKYKGVDGNOLPDSNRIDY	508	
D	b	455	NNLKASNTNKVDIASQVDDYNIKILGNSSTIIDDONTETIKKYVNSDOQLPQRSRIEDF	514	
O	y	509	SEVEDVTND-DYADLGNNNDVNINFGINDSPYLIIKIYSKSDPKKDYYTIIOGTVMOTTI	567	
D	b	515	SQYEDVTVSQEPNKKSFSSNNNAVTLDFGDINSAYLIKVSXKTPTPSDGELDIAGTSNRFT-	573	
O	y	568	NEYGEERTASDYDTITFFSTSOGOGDDL-PPEKTYKIGDIVWEVDVKGIONTNDEKP	626	
D	b	574	DKY-GYNYAKGISNEFYITSNDTDGGGDTGVAPBEKLKIDGYWVEDVDKRGVGOTBSKEPF	632	
O	y	627	LSNLVLTLYPPDGTGSVTRDEDGKYOFDLKANGLTYKTFPETREGYTPLKHSGTNPAL	686	
D	b	633	MANYLVLTLPDPGTTKSVFRPDANGHEFEGLJKGETEYTVKFEPTGYLPTKVNGTDDGK	692	
O	y	687	DSEENSWAVTINCGDDMTITDSGFYOPRKYSLGWYVVRDNKNKGIOQGDDKGI SGVVYVLTK	746	
D	b	693	DSNGSSTVYKINGKDMKSLDTGFYEKEYNLG DY WVEDTNKDG IO DANEPIKDV K VY L T K	752	
O	y	747	DENNISTSTT-----	757	
D	b	813	IKDADNMTLDGRGFYKTPKYSLSGYVWYDNSKNCKQDSTEKGIKIDVYTLQNKEGYIGTT	872	
O	y	758	-TDENGKYOEPDNLNSGWYIYHFDPKPSAMTOTTTDSCGDDDEQDADEEVAVITTDDHSFI	816	
D	b	873	KTDENGKYPFDNLDSGKYVIYEFKPAGLTQTVNTTEDD-KPADGGEVAVTITDHDDFTL	931	
O	y	817	DNGYDDESDDSD	876	
D	b	932	DNGFEEDT-----SD	980	
O	y	877	SD	936	
D	b	981	SD	1044	
O	y	937	SD	996	
D	b	1041	SD	1100	
O	y	997	SDSDSDSDSDSDSVSD	1054	
D	b	1101	SDSDSDSDGKHTPYKPM-----TTKDHNNKAALPE	1131	
O	y	1056	TGANEDYGSKCGLTGTLFACIGALL-GRKRNRK	1089	
D	b	1132	TGENNGSNNAITLFGILFALGSIULLFGRRKKQNK	1166	
RESULT 6 AAy83170 ID AAy83170 standard; Protein: 1802 AA.					
X	x	AAy83170:			
A	c	C	24-JUL-2000 (first entry)		
X	x	D			
X	x	E	Cell wall protein Sdrf.		
X	x	F			
R	w	Sdrf, SdrH: coagulase negative; staphylococcus; septicemia;			
K	w	osteomyelitis; endocarditis; immune response; vaccine; graft;			
R	w	sterile intravenous catheter; heart valve; cardiac.			
X	x	G	Staphylococcus sp.		
X	x	H			
K	e	Key Location/Qualifiers			
X	t	Misc-difference 12			









PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 XX Example 3; Seq ID No 13137; 511pp; English.

CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 1349 AA;

Query Match 36.6%; Score 2069; DB 22; Length 1349;  
 Best Local Similarity 36.6%; Pred. No. 1.6e-92;  
 Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;

QY 1 MIRKN-NLTKKKPIANKSNKYAIRKTVGTASIVGATLLFGHNEKAKEENSQDV 59  
 DB 1 MLNREKTAIRKGMVSNRKFSIRKTYGTASIVGATLLFGHNEKAKEESTKEL 60  
 QY 60 KOSNTDELSD 119  
 DB 61 NEAT-----SASDNOSD--KVDMOQIQEDNFKNDNQ--KEK-----VSSQGNSTTS 105  
 QY 120 STTNVENEATFLOKTPQDTHLTHEEVKSSSVESNSSIDPAQPSHTTINEEVSOT 179  
 DB 106 NGKKLLEKES--VOSTGKNVEVSTAKSDQASPKSTNEDLNTKQ-----TISQELAQ- 157  
 QY 180 SDVADSHVSDFAFNSKIKESSTESGKEENTIEQPNKYEDSTTSQPSGYTNI--DEKISNQ 238  
 DB 158 PDLQEKKSQVNV-----VQPTNEENKKVD-----AKTEST-----LNVKSDALKSN 198  
 QY 239 DELL--NLPINEYENKARPLSTTSADP-----SIKRVTVNOLAEOGSNVHLIKVTQSI 292  
 DB 199 DEFILVNNSSNNENNADIIILPKSTAPKRLNTRMRIRIAAVOPSPTEAKVNDLITSNTLT 258  
 QY 293 TEGYDSEGVIAKHADENLIVDTFEYDDKVKSGDTMTVIDKNTV-----PDLTDSF 346  
 DB 259 VVDADKNKKVTPADYLSLSQIT--VDDKVKSGDYFTIY--SDTVQVYGLNPDIKN-- 313  
 QY 347 TIRKID-NGEIIATGTYNKKNKQIYTTFDYDKENIKAHKLISYIDSKVPMNNT 405  
 DB 314 -IGDIDPNGEIIATKHTANLITFTDYDRNSQMGINSIYDADIP--VS 370  
 QY 406 KLDVEKTLASVNVKTIYEQRPN--ENRPAIQQSMFTNIDTKNHT-----VEQTI 455  
 DB 371 KNDVEFNVITGNTTKTANIQPDYVYVNEKNSIGSAFT--ETVSHVGNKENPQYQTI 428  
 QY 456 YINPLRASAETVNI-----SGNGDEGSTIIDSIIKVKVGNQNLPSNRIDYS-- 509  
 DB 429 YVAPSENSLTNAKLVQAVHSSYPNNIGQINKDVTIKITIQVPGYTL--HKGYDVNTK 485  
 QY 510 EYEDVNDYDAAQ--LGNNDVNIIFGNIDSPYIIKYSKYDPKNDYTTIQVTVQQT 566  
 DB 486 ELDDVNV-QYLQKITGDNNNSAVIDFENADSAYVWVNTKQYVNSSPILYQATISST 544  
 QY 567 INEYTGFRFASDNTASTSSGOGGDLPEKTYRIGDYWEDVDKDGIC-----618  
 DB 545 GN-----KSVSTGNALGFTNNOSGAG---DEVYKIGNYVWEDTKNGVQELGKGVG 594

QY 619 -----NTN-----621  
 DB 595 NNTVTFEDNNTNTRKVGAVTKEDGSYILPMLPBGDYRVESNLPKQEVTPSKGNNEL 624  
 QY 622 -----DN-----EKLNVLYTL-- 634  
 DB 655 DSNGLSSVTVNGKNDLNSADLGTYKPKYCNLDYWEDTNKNGIODEKIGSVTVTLKO 714  
 QY 635 -----TYPDG-----639  
 DB 715 ENGDIKTVTTADGKKYFTDLHNGNYKVEFTTPEGTPPTVSGSDIEKDSNGLATTGV 774  
 QY 640 -----TSKSV-----644  
 DB 775 INGADNNTLDSGFYKTPKYNLGNVWEDTKKDGKDSSTEGISGVYTLKNGEVLQTT 834  
 QY 645 RTDEGKRYOPDGLKNGLYTKITEETPEGYPPTLKSGTNPALDSEGS-----692  
 DB 835 KTDKDGKRYQFTGLENG--TYKVEFETPGYPTQVSGSTDEGIDSGNSTGTGVIKDKNDT 893  
 QY 693 -----VA-----VTI-----NGQ-- 700  
 DB 894 IDSGFYKPTYNLGDYWEDTNKNGVODKDEKIGISGVTVTLKENDKYLKTVTTDENCKYQ 953  
 QY 701 -----DDMTIDSGFYQTPK 714  
 DB 954 FTDLNNCTYVEETPEGYPPTSVTSGNDTEKDSNGLTGTVIKADNMTLDSGFYKTPK 1013  
 QY 715 YSLGNVWYDTPNKGIDGDEKIGISGVYTLKENDNGNIISTTTTDENGKYQFDNLNSGN 774  
 DB 1014 YSLGDYVWYDSNNDGKQDSTEGIKQVYKTVTLNEKKEGVIGTTRKDENGKYCFDNLDSGY 1073  
 QY 775 IVHFDKSGMTQTTTDSGDDDEDDADGEVAVYTTTDDPSIDNGYVDDSDSDSDSDSD 834  
 DB 1074 KYIFKEKPAGLQTVYNTTEDD--KDADGEVDVYTTDHDFTLDNGYFEEDT-----1123  
 QY 835 SDS 894  
 DB 1124 --SD 1181  
 QY 895 SDS 954  
 DB 1182 SDS 1241  
 QY 955 SDS 1014  
 DB 1242 SDS 1300  
 QY 1015 SDS 1073  
 DB 1301 -----TTKDHNNKAKALPETGSENNGSNNATLFGGLF 1332  
 QY 1074 AGLGALLL-GKRRKRRK 1089  
 DB 1333 AALGSLLLFGRKKQNK 1349

RESULT 10

AAU34082 ID AUU34082 standard; Protein; 932 AA.

AAU34082;

14-FEB-2002 (first entry)

Staphylococcus aureus cellular proliferation protein #358.

Antisense: prokaryotic cellular proliferation protein;

antibiotic; antibacterial; drug design.

Staphylococcus aureus.











Db 1074 KVFIEKPAFLQTGTNTTEEDKADADGGEVDVITTDHDFELDNGVYEETSDSDSDSDSD 1133  
QY 919 SPSDS 978  
Cc 1134 SPSDS 1193  
Db 979 SPSDS 1038  
Cc 1194 SPSDS 1253  
QY 1039 SPSDS 1086  
Db 1234 SPAGKHTPEKPMSTTKDHNKAKALPEIG-NENSGSNMNTLFGFLAALGSLILFGRKK 1312  
QY 1087 NRK 1089  
Db 1313 QNK 1315

RESURF 14  
U33975  
AAU33975 standard; Protein; 1021 AA.  
XX AAU33975;  
AC 14-FEB-2002 (first entry)  
DT 14-FEB-2002 (first entry)  
DE Staphylococcus aureus cellular proliferation protein #251.  
XX Antisense: prokaryotic cellular proliferation protein;  
KW antibiotic; antibacterial; drug design.  
OS Staphylococcus aureus.  
XX MO200170955-A2.  
PN 27-SEP-2001.  
XX 21-MAR-2001; 2001WO-US09180.  
PF 21-MAR-2001; 2000US-191078P.  
PR 23-MAY-2000; 2000US-206848P.  
XX 26-MAY-2000; 2000US-207272P.  
PR 23-OCT-2000; 2000US-242578P.  
XX 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
XX 16-FEB-2001; 2001US-269308P.  
XX (ELITR) ELITRA PHARM INC.  
A  
PI Haselbeck R, Ohlsen KT, Zyskind JW, Wall D, Trawick JC, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX MPI: 2001-611495/70.  
DR N-PSDB; AAS51834.  
XX  
XX New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids -  
XX  
XX Example 3; Seq ID No 5471; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the  
XX genes, their use in the discovery of novel antibiotics, the essential  
XX genes themselves and the encoded proteins. The prokaryotes used are  
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
XX invention is also useful for the identification of potential new targets  
XX for antibiotic development. The antisense nucleic acids can also be used  
XX to identify proteins used in proliferation, to express these proteins,  
XX and to obtain antibodies capable of binding to the expressed proteins.  
XX The proteins can be used to screen compounds in rational drug discovery  
XX programmes. The antisense nucleic acid sequence is also useful to screen

Cc for homologous nucleic acids which are required for cell proliferation in  
Cc a wide variety of organisms. The present sequence represents an  
Cc essential prokaryotic cellular proliferation protein.  
Cc Note: The sequence data for this patent did not form part  
Cc of the printed specification, but was obtained in electronic  
Cc format directly from WIPO at  
Cc ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 1021 AA;  
SQ  
Query Match 28.9%; Score 1634; DB 22; Length 1021;  
Best Local Similarity 37.4%; Pred. No. 1,6e-71;  
Matches 436; Conservative 155; Mismatches 334; Indels 240; Gaps 37;  
QY 18 KSNKYAIRKFTVGTASIVIGATLLEFG-LGHNEKAEENSVOYKDSNTDDELSDSDSS 76  
Db 5 KKEKAIRKKSIGVASVLGFLIGFLLSKKEADASENV-----TQSSASN 52  
QY 77 DEEKNDVINNNOSINTDNNQIIEETNNYDIEKRSDEPTESTTNVDENAEFTLOKTP 136  
Db 53 ESKSSDSSSVNAPKTDNTNVSPTKTSNTNNG-----ETSAQNPAQ--QETP 99  
QY 137 OD-NPHLEEEVK-----ESSSVSSNSIDTAQPSHTTINRESVQTSNVEDSHVSDFA 192  
Db 100 QASNTATTEETPVGEATTATKQANTPTQSSNT--NAELVQTSNETAS----- 151  
QY 193 NSKIESNTESGKEENTIQPNKVKEDSTTSQSGYTNIDEK--ISNDELNLPINEYE 250  
Db 152 NDTNTYSSVNSPQNSNNAENSVSTTODTSTEARPNSNESAPQSDANMKD-VYQAQAN--- 207  
QY 251 NKARPLSTTSAPQSIKRTVNOALAE--QGSNVNHLIKVTDOSTITEGYDSEGVKAKHD 307  
Db 208 -----TSA-PRMRAFSLAAVAADAPAGATDITN--OLTQVKT--IDSQTVYPHQ 253  
QY 308 AENLIYDVFEEVDKAKSGDPTMVDIDKNTVPSDLTDSFTPIKXNSGELITGYDK 367  
Db 254 AGYVKLNGFSVPNSAVKDDTEKITVPKELNLNGVSTAKVPIMYGD-QVLANGYDS- 311  
QY 368 NKQITFTDYDVKYENIKAKHLKLTYSIDSKVYNNNTKL-DVEYFTAL--SSVNTITY 424  
Db 312 DGNVITFTDYDVKENYANNTIMPAYID---PEAVTGTGNTLTGTGNTASTVLI 367  
QY 425 EYQRPNEENTANLQSMFTNIDTKNHTVEOTIYNPLRYSAKETNV-----NISNGDE 477  
Db 368 DYKQGFHNLISIKGTIDIDKNTNNTYROTIIYVP-----SGDNVLPALTLGMLINRTKS 422  
QY 478 GSFIIDSTTIIVYKVDONLPSNRIDYSEYEVDNDYQQLNNNDVNF-----G 533  
Db 423 NALIDAKNTDIAKIRV-DNANDISESYVNPSPDEVDYTNQVRISFNNANQKVEFTDDO 481  
QY 534 NIDSPYIIVISKYDPKNDYTTIQOTVTMTQTTINEYTGEE--RTASDYNTIAFTSSGQ 591  
Db 482 QITTPYIVVNGHIDPASNG-----DLALRSFFGYDSNFIKRSMSWNEVAFNNGSGS 535  
QY 592 GCG-DLP--PEKTYKIGDY--VWEDVDKGIQONTNDNEKPLSNVLYTLTPDGTSASVPT 646  
Db 536 CGGIDKPVPEQDEPEGEIEPIPEDSDSD-----PGSDSGSPSN 574  
QY 647 DEDGKQFGDLNGLLKYKTFETPEEGYPTLKHSGTNPALDSGNSVWYITNQD-DMT 705  
Db 575 SDSG-----SDSGSDSTSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 607  
QY 706 DSGFYQTPYSLGNYVYDNTKNGIGDEKIGISGKVLKDENGNINISFTTDEMGKYQ 765  
Db 608 DS-----DSTSDSDASDSASD-----SDSASDSASD 638  
QY 766 FDLNLSGNYIVHDKPSGMTQTTSDGDEDDADAEVHVTTTDDHDFSIDNGYDDES 825  
Db 639 SPSASDSDSDSDSDS--SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 697  
QY 826 DSD 885  
Db 698 DSD 757





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01; Search time 17.6129 Seconds

(without alignments)  
1824.221 Million cell updates/sec

Title: US-09-147-405b-15

Perfect score: 5646  
Sequence: 1 MINKNNLLTKRRKPIANKSN.....FAGLCALLGKRRKRRKKN 1092

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCROS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4490	79.5	930	4	US-09-134-001C-5314
2	1628	28.8	933	3	US-08-293-728-2
3	1628	28.8	933	4	US-09-421-868-2
4	635.5	11.3	2137	4	US-09-134-001C-4463
5	482	8.5	466	4	US-09-134-001C-4749
6	377.5	6.7	1183	4	US-09-134-001C-3530
7	349	6.2	345	4	US-08-856-253-7
8	331.5	5.9	1161	4	US-09-327-536-2
9	315.5	5.3	1112	2	US-08-714-402-2
10	301.5	5.3	1231	3	US-08-904-263A-4
11	286	5.1	1060	4	US-08-911-393-2
12	282	5.0	3052	2	US-08-557-122A-26
13	282	5.0	3052	4	US-09-162-666-26
14	279	4.9	3696	4	US-09-134-001C-5080
15	278	4.9	699	4	US-09-134-001C-4054
16	276.5	4.9	1085	1	US-08-431-080-28
17	276.5	4.9	1085	2	US-08-938-534-28
18	276.5	4.9	1085	4	US-09-345-294-28
19	276	4.9	2391	2	US-08-446-855A-2
20	276	4.9	2391	4	US-09-150-741-2
21	271	4.8	1115	2	US-08-568-459A-2
22	271	4.8	1115	2	US-08-487-826B-2
23	271	4.8	1115	2	US-08-487-826B-2
24	271	4.8	1115	6	5198347-6
25	270.5	4.8	1833	4	US-08-621-944A-4
26	270.5	4.8	1833	4	US-08-621-944A-4
27	270.5	4.8	1992	4	US-08-621-944A-3

28	270.5	4.8	1992	4	US-08-945-567D-3	Sequence 3, Appl1
29	269.5	4.8	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
30	267	4.7	2123	4	US-08-968-685A-10	Sequence 10, Appl
31	265	4.7	2048	4	US-09-268-347-48	Sequence 48, Appl
32	263.5	4.7	2314	4	US-09-268-347-49	Sequence 49, Appl
33	260	4.6	1183	2	US-08-447-031A-2	Sequence 2, Appl1
34	259.5	4.6	1435	2	US-08-568-459A-4	Sequence 4, Appl1
35	259.5	4.6	1435	2	US-08-487-826B-4	Sequence 4, Appl1
36	259.5	4.6	1435	2	US-09-210-288-4	Sequence 4, Appl1
37	257	4.6	1776	4	US-09-556-877-179	Sequence 179, Appl
38	257	4.6	1776	4	US-09-556-877-179	Sequence 179, Appl
39	252.5	4.5	1786	4	US-08-973-462-8	Sequence 8, Appl1
40	252	4.5	414	1	US-07-667-276A-4	Sequence 4, Appl1
41	250.5	4.4	821	4	US-09-556-877-195	Sequence 195, App
42	250.5	4.4	821	4	US-09-620-412C-195	Sequence 195, App
43	250	4.4	690	4	US-09-134-001C-4568	Sequence 4568, Ap
44	249.5	4.4	905	4	US-09-074-658-70	Sequence 70, Appl
45	245	4.3	1850	4	US-09-620-093A-5	Sequence 5, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-134-001C-5314  
Sequence 5314, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5314  
LENGTH: 930  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5314

Query Match	79.5%	Score 4490;	DB 4;	Length 930;
Best Local Similarity	81.2%	Pred. No. 1.8e-241;		
Matches 884;	Conservative 19;	Mismatches 26;	Indels 160;	Gaps 3;
QY	4	KNNLLTKRRKPIANKSNKVAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVQDYKDSN 63		
DB	2	KNNLLTKRRKPIANKSNKVAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVQDYKDSN 61		
QY	64	TDELDSDNOSDEERNDYNNNSQINTDDNQIITKEETNYVDGIEKESDEPTESTN 123		
DB	62	MDDELDSDNOSDEERNDYNNNSQINTDDNQIITKEETNSDNLAKESNKTITOSTN 120		
QY	124	VDNEATFLOKTDQDNTHLFEFEVKESSESVSSSIDFQOQPSHTTINREESVO 183		
DB	121	VDNEATFLOKTDQDNTHLFEFEVKESSESVSSSIDFQOQPSHTTINREESVO 180		
QY	184	EDSHVDSFANSKIKESNTESGKEENTIEOPNKYEDSTTSQPSGYTIDEXISNOBELN 243		
DB	181	ENSRVDSFANSKIKESNTESGKEENTIEOPNKYEDSTTSQPSGYTIDEXISNOBELN 240		
QY	244	LPINEYENKRPSTTSQAOSIKRVYVNOAAGSVNHLIVTDOSTIEGDDSGV 303		
DB	241	LPINEYENKRPSTTSQAOSIKRVYVNOAAGSVNHLIVTDOSTIEGDDSGV 300		
QY	304	KADAENLIVDYFEVDKDKKSGDTMTVDIDKNTVPSDLTDSFTIPKIKNSGEIATGT 363		
DB	301	KADAENLIVDYFEVDKDKKSGDTMTVNDKNTVPSDLTDSFTIPKIKNSGEIATGT 360		







[illegible]

```

: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 4749
: LENGTH: 466
: TYPE: PR1
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4749

```

Query Match	8.5%;	Score 483;	DB 4;	Length 466;
Best Local Similarity	52.1%;	Pred. No. 1,4e-19;		
Matches 113;	Conservative 25;	Mismatches 53;	Indels 26;	Gaps 4

QY 746 KDENGNIITTDENCKYQFDNLNSGNIIVHFKDKPSGTQTOTDSDGDDEADAGEVH 805  
| | | : : : : :  
Db 9 KDCNGEVV-----EEMKKFN-IKHSFMLTGFAFMYTTSLSFSQAHHAG----- 50

Db 51 -----NHPIIDINFQDIDRNTAKSNILNRKNDTSRIGISMNSDNDLDTDIYNSDSE 103

[illegible]

	Gy	925	SDDSDSDSDSDSDSDSDSDSDSDSDS	961
			:	
Pd	164	SDDSDSDSDSDSDSDSDSDSGTGGKS	200	

RESULT 6  
US-09-134-001C-3530  
Sequence 3530 Amplification ITS/09134001C

Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: LYON DOUCETTE-STAMM ET AL  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPH

EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964  
 PRIOR FILING DATE: 1997-11-08  
 PRIOR APPLICATION NUMBER: US 60/055,779  
 PRIOR FILING DATE: 1997-08-14

```

: NUMBER OF SEQ ID NOS: 56/4
: SEQ ID NO 3530
: LENGTH: 1183
: TYPE: prt

```

```

Organism:Staphylococcus epidermidis
US-09-134-001C-3530
Querry Match 6.7% Score 377.5 DB 4. Length 1183.

```

2 INKKNNITKKKEDIANSSNKVARIKPEKTVGTAATVITGATITLFGHGNHNAK-AAPENSVDVKK 60  
Matches 264; Conservative 199; Mismatches 495; Indels 343; Caps 56

```

15 INKKVDEL-----PKNLKYSIRKFTVGASTILGGSTLIFGSSSHEAKAAEEEOVDPIIT 68
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
61 DNSTDELSDSNQOOSDDEKENVYTN-----NOSTINFD-----NNOTIKKRETNND 108

```

[illegible]

```

      125 APOSEPTKAEEGNAEAPOSEPTKAEEGNAEAPOSEPTKAEEGNAEAPOSEPTKTEEG 184
      151 SSVES-----SNSSTDTAAO-----DGHMTTARDFSV----- 177

```

Db 185 SNAEAPNVPTIKANSNDNTOTOFSEAPTRNDLARKEDIPAVSKNEELOSSQPNPTDSKIEP 244



Db 310 IWRSMWMDNEVAFNNNGSGGDIDKPVVPEQ 340

## RESULT 8

US-09-327-536-2

Sequence 2, Application US/09327536  
Patent No. 635477

GENERAL INFORMATION:

APPLICANT: FISCHETTI, Vincent A.

APPLICANT: ROCHA, Claudia

TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A

TITLE OF INVENTION: STREPTOCOCCI

FILE REFERENCE: 02927-008

CURRENT APPLICATION NUMBER: US/09/327,536

CURRENT FILING DATE: 1999-06-08

PRIOR APPLICATION NUMBER: US 08/714,402

PRIOR FILING DATE: 1996-09-16

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 1161

TYPE: PRT

ORGANISM: SFBBP gene

US-09-327-536-2

Query Match 5.9%; Score 331.5; DB 4; Length 1161;  
Best Local Similarity 18.8%; Pred. No. 9.9e-11;  
Matches 244; Conservative 183; Mismatches 447; Indels 421; Gaps 54;

Db 18 KSNKYAIRKFTVGTASIVYCATLFP-----GLGHNEAKAEENSVQDYKDSNTDELSDS 71  
Db 3 QKNSYKTL-SPLSLTLGFLGLLVLFIGLSGVSGHAETR-----NGA 43  
Qy 72 NDSSDEKKNNVNNNSINFDNNQIIRKEETNNYDGIKRSRDESTTNVDENATF 131  
Db 44 NKQSSFE-----IKKVDQNN----- 58  
Qy 132 LQKTPDQNTLHTEEBKES--SVESNSSIDTAQ--OPSHITINRESVQTSQNVEDSH 187  
Db 59 -KPLPGATSLTSKDGKTSVQSFSTNSDKGLVDQNLQPGTYTLKEPAAPGVC----- 111  
Qy 188 VSDRANSKIKSNTESGKEENTTEOPNKVKEDSTTSQPSGTTNDEKISNDELINLPIN 247  
Db 112 -----KTSRTW-----TVVYENGTYTKLVENPN-----G 136  
Qy 248 EYENKARPLSTSAQPSIKRVTYNQLAEGSNVNHLLKVDQSTEGYDSEGVYKAHD 307  
Db 137 EIRKAGSKDVSSSL-----QLENPKMSVVSKEYG-----TEVSSGAAPFYRNA 181  
Qy 308 AENLIYDTEFVDDKVS-----GDTMTVDIDKNTVPSDLTDSFTIPK-IKNGCEIAT 361  
Db 182 A---YFKMSFELKQKDKSETINPGDTFVLQDRRLNPKGISOD--IPKIITDSANSPLA 236  
Qy 362 GTYNKKNQIYTFDYYDKYENIKAHKLKLSYIDSKVPPNNMTKLDEYKTAASSVNT 421  
Db 237 GKHYAENHQLYTFDYYAGDKVQLSAELSLPLENKEVLEN-----TJSINFSST 287  
Qy 422 I-----TVEQRPNE-----NRANLQSMPTNIDTKNHYEQIYIINPLAYS 463  
Db 288 IGGGEITYKGTGNVLYGNESTKESNYITNGLSNGSGSIESNTETGEFVWVYVNPNTN 347  
Qy 464 AKETNVNINSGN-----DEGSTIIDSTIIKVKYVDQNDQNLPSDNRIYDSEVED 513  
Db 348 IPYATMNLMGGRARSNTSDLENDANTSALGELYVEVEGKELPSSYGV-----D 400  
Qy 514 VT-----NDYIAQGNNDV-----NINEGN--IDSPYIIKVISKYDPKKDDYTTIQQT 562  
Db 401 VTKTLRLRIDTAGLNGFGMTKQRIDEGNNIYQNKAFIITKGTGTDOS-----GRKLV 453  
Qy 563 MOTININEYGEFRAS-----YDNTIAFSNSSGGGCG-----DLPEKTYKIGD 606  
Db 454 VQSNLASFRGASEYTAFTPVGGNVYFQNEIALSPKSGSGSGKSEFTKPSITVANLKRVAQ 513

Qy 607 YWEDVDKGIQNTNDNEKPLSNLVLTLYPDGTSKSVR--TDDEGKYQFDGLKNGLYTK 664  
Db 514 LRFKMSSTDNY-----PLPEAFELRSNGNSQKLEASNTQGEVHFKDLTSG--TYD 564  
Qy 665 ITFET--PEGI-----TPTLKHSQTNPA-----LD 687  
Db 565 L-YETKAPKGYQVTEKRLATVTDTRPAEEMVYWGSPHSSVKVEANKVPIVNHKELTL 623  
Qy 688 SEGNSVW-----VTINGO-----DDMTIDSG-----FYQTRKYS----- 716  
Db 624 FSGKIMENDRPDRPAKIYQVLQNGQKMNQIQEVTAKDMWSYHFKDLFKYDAKNEY 683  
Qy 717 -----LGANYW-----YDINKDGIQ-GDDE-KGISGVK-----TL 745  
Db 684 KYSEEVNVPDGYKVSYLGNDFPTRETEFEFQDNFNFLERGNMEIKQSGSKIIDETL 743  
Qy 746 KDENGNIISTTTDEN-----GKIQFQNL-----NSG 772  
Db 744 TSFGKRIKWKNDTAENRQAIQVQLYADGVAEGQTRISGSGNEMWSEFENLKRKYNCTG 803  
Qy 773 NYIVHFDKPSGMTQTDTSDGDDQDADAGEVHTYTPDHDFSIQNGYYDDESDSDSDSD 832  
Db 804 NDIY-----SVKEYTVPYTGIVTYSAN--DIINTKREVIYIOGPKLEIEETLPLESG 854  
Qy 833 SDSDSDSDSDSD--SD 885  
Db 855 ASGGTTFVEDSRPDTLSGLSDEGQSGDMTEEDSATIHFKSRKIDGKELAGATWMLR 914  
Qy 886 DSDSDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 931  
Db 915 DSSGKTITWTISDQVKDFYLPKRYTFEETAPDGYEIAIAIFTVVEQGVYVNGKAT 974  
Qy 932 DSDS-----DSD 977  
Db 975 KGDTHYVWDAKPKTKGSGQYIDIEKLPDQHGSGSTTEIEDKSSDLIIIGGQEVVD 1034  
Qy 978 DSD 1031  
Db 1035 TEDQSGMTGSHSGSTTEIEDKSSDVIIGGQGVYETEDPOTGYGDSGCKTEVENK 1094  
Qy 1032 -----DSD 1059  
Db 1095 VQSFHFDNKEPESSNEIIPKDKPKSNTSLPATGEN 1129

## RESULT 9

US-08-714-402-2

Sequence 2, Application US/08714402

Patent No. 5910441

GENERAL INFORMATION:

APPLICANT: FISCHETTI, Vincent A.

TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING

TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHEIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,402

FILING DATE: 16-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McGowan, Malcolm K.  
 REGISTRATION NUMBER: 39,300  
 REFERENCE/DOCKET NUMBER: 016921-097  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1112 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-714-402-2

Query Match 5.6%; Score 315.5; DB 2; Length 1112;  
 Best Local Similarity 20.1%; Pred. No. 7.2e-10;  
 Matches 181; Conservative 137; Mismatches 279; Indels 303; Gaps 43;

18 KSNKVAIRKFTVGTASIVIGATLLE-----GLGHNEAKAEENSVDVKDSDNDELSDS 71  
 3 OKNSYKL-SELLSLTGFIILLGLVIGLVSGHAETR-----NGA 43  
 72 NQSSDSEKNDVYINNOSITDDNNOIKKEETNNYIGIKRSDEKREESTNVENRATF 131  
 44 NKGSFE-----IKKVDNN----- 58  
 132 LQKPDNTHLREEVEKSS--SVESNSSIDTAQ--QPSHTINREESVQTSNVEDSH 187  
 59 -KLPKATSLTSGDKDGTSGVOSTSNDKGIYDAONIQPGTYTLKETATADGVP----- 111  
 188 VSDPANSKIKESNTESGKEENTIEQPNKVKEDSTTSQSGVTNIDEKISNODELLNLPIN 247  
 112 -----KTSRTW-----FTVYENGTYKLV-----N 132  
 248 EYENKARPLSTSAQPSIKRYVNOALAEQGS-NVNHILIVTD--QGITEGYDSE--G 301  
 133 PYN-----GEIISKAGSKDVSSSLQLEENKMSVSKYGGTEVSSG 172  
 302 VIKAHDAENLIYDVFEDDKVKS-----GDTMTVDIDKMTVPDSLDFSTPK-IKDNS 355  
 173 AADFYRHHAYFKMSFELKDKDSETINPDFTYLOLDRINPFGICOD--IPRIITDSA 230  
 356 GEIATGTYDNKKKQITTYFTDYVYKXENIKAKHLKLSYIDKSKVPPNMTKLDEYKTA 415  
 231 NSPLAIGKYHAENHQLITFTDYIAGLDKQVLSAELSLFLENKRYLEN-----TSI 281  
 416 SSYVKTI-----TVEQORENE-----NRTANLQSMFTNIDTKNHYVEQTIYI 457  
 282 SNKSTIGGGEITYKGVNVLVYGNESKRESNYITNGLSNNGSGSIESYNTETGEFVWYVY 341  
 458 NPLRYSAKETNNVNISSNG-----DEGSTIIDSTIIKVKYKVDNQLPNSRIYD 507  
 342 NPKRTINIPYATMLMGFGARSRNMTSLENDANTSSAELGETOYVEVEGEKLESYGV-- 399  
 508 YSEYEDVT-----NDVYALQNNNDV--NINFGN--IDSPYLIKYSKYDPKRDYTT 556  
 400 -----DVTKLTLRTDITAGNGFQMTKRORIDFGNNIQAFLIKYTKGTDOS----- 448  
 557 IQQVTVMQTTINETGEFFRAS-----YDNTIAFTSSGGGOG-----DLPEPK 600  
 449 -GKPLVVOSSMLASFRGSEYAAFTPVGANYVFOINEILALSPKSGSGSKSEFTKPSIVAN 507  
 601 TYKIGDYVWEDVDKGIQNTNDNEKPLSNVLVLTYPDGRSKSVR--TDDEKAKQFQGLK 658  
 508 LKRVAAQLRKKMSTDVN-----PLPEAFAELRRSSNGNSOKLEASSNFGVEHFRLDT 559  
 659 NGILTYITFEET--PEGY-----TPTLKHSGTNPALDSEGSNVY-----VTINGQDDMT 704  
 560 SG-TYDL-VETTKAPKGGQVTEKLAITYVDTPTKRA--EEMVYMGSSHSYKVEANNEVT 614  
 705 ISGSGYQTPKYSIGNYVWYTNKDGIGQDDEKIGSVKVTLKDENG--NISTTTTDE 760

DB 615 I-VNHKETTAFS-GKKIWMENRDP-----QREPAKIQVOLL-ONGKMPNQIQEVTKDN 664  
 QY 761 NGKYQFONLNSGANYIVHFKRPSGWTQTDTSDGDEDDADQEEVHVITTHDDDSIONGY 820  
 DB 665 DMSYHFKDL-----PKTDAKQOETKYYSV--EEVNVDPGX 696

# RESULT 10

US-08-904-263A-4  
 Sequence 4, Application US/08904263A  
 Patent No. 6015889  
 GENERAL INFORMATION:  
 APPLICANT: LINDAHL, GUNNAR  
 APPLICANT: STALHAMMAR-CARLEWALM,  
 APPLICANT: MARGARETHA  
 APPLICANT: STENBERG, LARS  
 TITLE OF INVENTION: PROTEIN RIB, A CELL SURFACE PROTEIN THAT  
 TITLE OF INVENTION: CONTERS IMMUNITY TO MANY STRAINS OF THE GROUP B  
 TITLE OF INVENTION: STREPTOCOCCUS: PROCESS FOR PURIFICATION OF THE PROTEIN,  
 TITLE OF INVENTION: REAGENT KIT AND PHARMACEUTICAL COMPOSITION  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP  
 STREET: PO BOX 747  
 CITY: FALLS CHURCH  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22040-0747

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/904,263A  
 FILING DATE:  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: SVENSSON, LEONARD R.  
 REGISTRATION NUMBER: 30,330  
 REFERENCE/DOCKET NUMBER: 552-119P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 205-8000  
 TELEFAX: (703) 205-8050  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1231 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-904-263A-4

Query Match 5.3%; Score 301.5; DB 3; Length 1231;  
 Best Local Similarity 19.0%; Pred. No. 4.9e-09;  
 Matches 258; Conservative 198; Mismatches 494; Indels 409; Gaps 62;

3 NKKNNLTKKKKPIANKNKVAIRKFTVGTASIVIGATLLR----- 43  
 8 NSIDTLQTKQ-----RESIKKFKFGASVLIGISFLGFGTGGOFNISTDTVFAAEVI 59  
 44 -----LGHNEAKAEEN-----SVQDVKDSNTD----- 65  
 60 SGSAVTLNTMNTKRVQNGRAYIDLVDKNGKIDPDLQILTLNSPLKQAYIROGNGFTQ 119  
 120 PSELTIVGAASINTVLTGDSGPHTKPDGVDIINSLTIYNSALRDKIDEVKKRAEDP 179  
 106 NYGIEKRSFDR-----ESTNVNEN-----EATFLQK-----TPDNT 140  
 180 KWD--EGSRDKYLIISDDIKTIDNNPKTQSDIANRITETVNTLEKILVRIPADKNDP 236  
 141 HLTREEVEKSSSVESNSSIDTAQPSHTTINREESVQTSNVEDSHVSDPANSKIKESN 200



```

Db 256 SDFGSEPALEMEGLIINTFVDRRLVYESSDNGSMVEAVGLSR--VMGSPKS 312
Qy 432 NRTANLOSMTNIDRKNTVEQTIYINPLRYSAK-----ETNVISGNGEGSTIIDSTI 487
Db 313 NQPGS-QSSEFTAVTIEGMKV--MLFTHPLNFKGRWLRLNLTMLTNO-----357
Qy 488 IVYKVGDNQNLPSNRIVDYSEYEDVTNDYAQLGNNDVNINFGN-IDSPYIIKVISK 546
Db 358 -RIYVNG-QVSIIGDENSAVSVLYKDKDLCLYCLHEINSNEYSLVFAFLVGLALIKSVLD 415
Qy 547 YDPNKDDY-----TTIOQT--VTMOTITNEYGER-----T 576
Db 416 SKKMNDSHLSICTPADPAASSERCCGPAVTVGLVGLSHSATTEWEDATRCVNAST 475
Qy 577 ASYD---NTIAFTSSGOGGDLPP-----EKTKIGDVMEVDYDKDGIONTNDNEKPL 627
Db 476 ANAERPNLKF---AGVGGALMPYSOGGQONRHFANHAFLVASVTHHEPVSASPL 552
Qy 628 SNVLYTLTYPDGTSKVRTEDEGKYQFDGL-----KNGLYTKITFE-----668
Db 533 LGASLD---SSGKLLGLSYDEKHOMQPIYGSTPVTPTGSMEMGRKRYHVLTMANKIGS 589
Qy 669 -----TPEGYPTLKHSCTNPALDSEGNVWYVINGODDQITIOSGFYQTP 713
Db 590 VYIDGPELGSGQTVVDPDRTPDISH-----FYVGGRSDHPITIS-----630
Qy 714 KXSLGVYVYDTKDKGIDGDD-----EKISGVKYLKDEKNITIST--TTDENGKYQ 765
Db 631 HTYVNVLYLXNRO---LNAEITLFLSODDLITEAHMGSSSSASHSTISTADNCAHS 687
Qy 766 FNULNSGNYIVHDKPSGNTQTT-TDSGDDDEQDADGEE-VHVTITDHDFFS-----815
Db 688 TPSTPA-----DSASHSTPSPADSAHSTPSPADNCAHSTPSPGDSASHSTPSTP 740
Qy 816 IDNGYVDEDS---DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 872
Db 741 ADNGAHTPSPADSAHSTPSPADNCAHSTPSPADNCAHSTPSPGNGAHTPSTP 800
Qy 873 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 932
Db 801 GDSASHSTPSPADNCAHSTPSPADNCAHSTPSPADNCAHSTPSPADNCAHSTPSTP 860
Qy 933 SPDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 992
Db 861 ADSASHSTPSPADNCAHSTPSPADNCAHSTPSPADNCAHSTPSPADNCAHSTPSTP 920
Qy 993 SDSDSDSDSDSDSDSDSDSDSDS-----DSGSDSDSDSDSDSD-----1036
Db 921 GDSNAHSTPSPADSAHSTPSPADSAHSTPSPADNCAHSTPSPADSAHSTPSTP 980
Qy 1037 NPSDL-----GNSSDKSTKDKLPTPGANE-----DYSGKGTLL-----G 1070
Db 981 GDSASHSTPSPADSAHSTPSPADNCAHSTPSPADNCAHSTPSPADNCAHSTPSTP 1040
Qy 1071 TLFAGLALL 1081
Db 1041 GLLLCAGALL 1051

```

```

; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambitis, Elias J.
; REGISTRATION NUMBER: 33,728
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-557-122A-26

Query Match 5.0%; Score 282; DB 2; Length 3052;
Best local Similarity 19.7%; Pred. No. 1,9e-07;
Matches 207; Conservative 147; Mismatches 418; Indels 280; Gaps 47;

Qy 141 HLTEEEYKSSSVSSNSSIDTAQPSHTTINREESVQTSNDVSHVDFAN-----SKI 196
Db 2129 HLMSQELPDE-----DMDRQPVKVLVGRKPE---SDDLTADIKIFCKNKTLEGI 2175
Qy 197 KESNTESGKEENTIEOPNKKEDSTSQPSGYNIDEKISNDEL---LMLPINEYENKA 253
Db 2176 KPHLSQDLPEDDMDKQPVK-----LVGKDELSDKIVLESKAIESLB 2217
Qy 254 RPLSTTSQPSIKRVY-----NOLAEQGSNV-----NHLI 285
Db 2218 KDFLKGDASIVASQELFEFENODSSVFLVQKNFEVAFDEKKNVFEVYAPMGCHQOLA 2277
Qy 286 KYVDOSITGEYDSEGVYKADAENLYDYTFEYVDKVKSGDTMTYDID-----KN 336
Db 2278 PIMDK-LGETYKDHENIIIAKNFEVAFD-----EKKNVFEVYAPMGCHQKQ 2324
Qy 337 TVPSDLTDSFTIRKIKNSEIITATGY-DNKKQIYTTFTD-----YDKYENIK 366
Db 2325 LAP-----IMDKLGE---TYKDHENIIIAKNFEDVAFDEKKNVFEVYAPMC 2368
Qy 387 AHLK-----LTSYIDSKSVPPNNNTRKLDVEYKTAALSSVKNKTIYEXORP-----429
Db 2369 GHCKQLAPIMDKLGETYKDHENIIIAKNFEVAFDEKKNVFEVYAPMGCHQOLAPIMD 2428
Qy 430 -----NENRTANLOSMTNI---DTKNHTEQTIYINPLRYSAK-----466
Db 2429 KLGETYKQHODIYIAKNFEVAFDENKNVFEV---FYAPMGCHQOLAPAMDKLGPTYRD 2485
Qy 467 -TNVNISGNGDEGSTIID--STIIKYKV---GDONLPLDSNRIYVSEYEDVTNDYA 520
Db 2486 HENIVIAKNHDE--IYNDKDKDLVLYVAPMGCHCKRLAPT-----YQELA-DTYA 2533
Qy 521 OLGNNDVNINFGNIDSP---YIIKVIS---KYDENKDDYTTIQ-----558
Db 2534 -----NATSVDLIAKMDSTANEVAVKHSPTLKFPPASADRYVIDNGERTLDGFKFL 2589
Qy 559 QTVMTQTTINEY-----TGEFRTASVDMNTIAFTSSGGQGGDL-PEPKYKISDYVW 609
Db 2590 ESGGMDSTANEVAVKHSPTLKFPPASADRYV-----DYNGERLT---DGFK 2636
Qy 610 EDVDKDGIOQT-NDNEKPLSNVLYTLTYPDGTSKSVTDEDEGKYQFGLKNGLYTKITFE 668
Db 2637 KFLSSGMDSTANEVAVKHSPTLKFPPASADRYVIDNGERTLDGFKFL-----2689

```

RESULT 12  
 US-08-557-122A-26  
 ; Sequence 26, Application US/08557122A  
 ; Patent No. 5879664

## GENERAL INFORMATION:

APPLICANT: Hjort, Carsten Malland  
 TITLE OF INVENTION: Fungal Protein Disulfide Isomerase  
 NUMBER OF SEQUENCES: 38  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5879664 of No. 5879664 disk of No. 5879664th America, Inc.  
 STREET: 405 Lexington Avenue, 64th Floor

CITY: New York  
 STATE: New York  
 COUNTRY: United States of America

```

QY 669 TPEGTPPLKHSCTPALDSEGSVWWTINGODMTIDSGFYQTRKYSIGANTVWDTNKD 728
Db 2690 -----ESG---GMDSTANEV-----EAVKVS--FPTLKFFPAG-----PGRT 2722
QY 729 GLOGDEKISGVKTVLKDENGNIITTTDENGK-YQPDNL-----NSGNTIYHDKP- 781
Db 2723 VIDYGERLIDGFKFL-ESGGMSTANEVAVKVSFPTLKFFPAGSGRNVIDNGER 2780
QY 782 --SGMTQTTTDSG-DDDEODADGEEVHTITDHDFFSIDNGYDDESDDSDSDS-DSDS 837
Db 2781 TLEGKFKFLESGLDHTENDVARG---VYIEGYPTIVLYPGKKSSSVYQGSRSLSLFL 2836
QY 838 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 897
Db 2837 DFIKENGGDAGDND-DLDLEALPDMEDDDOKAVKDELQDAGDDDLLELEAEER 2895
QY 898 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 957
Db 2896 DLEEDDDOKAVDELQDAGDDDLLELEAEERPMEDDDOKAVKDELQDAGDEGL 2955
QY 958 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1017
Db 2956 DLEAEERPOLLEEDDDOKAVRDELQDGAADDDLELDETEERDLE-EGDDDEOKIKQKDEL 3014
QY 1018 DSDSDSGS---DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1046
Db 3015 HEDVDGKALYEAEOKAEAEADADADELAEED 3046

RESULT 13
US-09-262-666-26
; Sequence 26, Application US/09262666
; Patent No. 6346244
; GENERAL INFORMATION:
; APPLICANT: HORT, Carsten Maitland
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 63462440 No. 6346244disk of No. 6346244th America, Inc.
; STREET: 405 Lexington Avenue, 64th floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,122
; FILING DATE: 11-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 380,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-262-666-26

```

```

Query Match          5.0%; Score 282; DB 4; Length 3052;
Best Local Similarity 19.7%; Pred. No. 1.9e-07;
Matches 207; Conservative 147; Mismatches 418; Indels 280; Gaps 47;

QY 141 HLTREEVSSVSSSSSIDPAQPSHTTIRRESVQTSVDNEVSHVDFAN-----SKI 196
Db 2129 HLMSELPDE-----DMDROPVAVLVGKKE---SDDLTKDKIKEFNKFLGKTI 2175
QY 197 RESNTEGKEENTIEQPNKVKEDSTTSQPSGYTNIDEXISNDEL---LNLPINEYENKA 253
Db 2176 KPHLMSQLPEWMDKOPKV-----LVGKDELSDKIVLESKAIESJB 2217
QY 254 RPLSTTSAPSTKRVTV-----NOLAEOGSNV-----NHLI 285
Db 2218 KDFLKGDASPIYKSGEIEFENDSSVFOLVGNFEVEAEDEKKNVFEVEYAPMGCHCKOLA 2277
QY 286 KVTDSITEGYDSEGVKAKHDAENLIYDVPEYDOKVSGGTMVVDI-----KN 336
Db 2278 PIWDR-LGETYKDHENIILAKNFEVARD-----EKNVVEFYAPMGCHCKQ 2324
QY 337 TVPSDLTDSFTIPKIKNSGELIATGY-DNKKOITYFTD-----YDKYENIK 386
Db 2325 LAP-----IMDKIGE---TYKDHENIILAKNFEVARDKKNVFEVEYAPMC 2368
QY 387 AHLK-----LTSYDKSKVPNNNTKLDVEYKTALSSVNTIYVYORP----- 429
Db 2369 GHCKOLAPIMDKIETGYKDHEIILAKNFEVEAEDEKKNVFEVEYAPMGCHCKOLAPIWD 2428
QY 430 -----NENRANTLOSMFTNI---DTKNHTVEQTIYINPLRSASE----- 466
Db 2429 KLGETYKEHODIYIAKNFEVEAEFDEKNVFE---FYAPMGCHCKOLAPAMKLGPIYRD 2485
QY 467 -TNVNISSGDEGSTIIDD-STIIRYKV---GNQNLPSDNRIYDYSEYEDVNTDXYA 520
Db 2486 HENIYIAKNHDE---YVNDPKRDVLVLYAPMGCHCKRLAP-----YQELA-DTYA 2533
QY 521 QLGNNNDVAINFGNIDSP---YIIRKVIS---KDPKNDDYTTIQ----- 558
Db 2534 ---NATSDVLIAMKDSANEVAVKVSFPTLKFFPASADRTVIDYNGERTLDGFKFL 2589
QY 559 QTVTQMTITNEY-----TGEFRTASYDNTIAFSTSSGOGDLPPEKTYKIGDIYW 609
Db 2590 ESGGMSTANEVAVKVSFPTLKFFPASADRTVI-----DYNERL---DGK 2636
QY 610 EDVYDKGIQNT-NDNEKPLSNVLVTLPDGTSKSVFDEDEGKYOPDGLKNGLYTKIYFE 668
Db 2637 KLESGGMSTANEVAVKVSFPTLKFFPASADRTVIDYNGERTLDGFKFL----- 2689
QY 669 TPEGTPPLKHSCTPALDSEGSVWWTINGODMTIDSGFYQTRKYSIGANTVWDTNKD 728
Db 2690 -----ESG---GMDSTANEV-----EAVKVS--FPTLKFFPAG-----PGRT 2722
QY 729 GLOGDEKISGVKTVLKDENGNIITTTDENGK-YQPDNL-----NSGNTIYHDKP- 781
Db 2723 VIDYGERLIDGFKFL-ESGGMSTANEVAVKVSFPTLKFFPAGSGRNVIDNGER 2780
QY 782 --SGMTQTTTDSG-DDDEODADGEEVHTITDHDFFSIDNGYDDESDDSDSDS-DSDS 837
Db 2781 TLEGKFKFLESGLDHTENDVARG---VYIEGYPTIVLYPGKKSSSVYQGSRSLSLFL 2836
QY 838 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 897
Db 2837 DFIKENGGDAGDND-DLDLEALPDMEDDDOKAVKDELQDAGDDDLLELEAEER 2895
QY 898 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 957
Db 2896 DLEEDDDOKAVDELQDAGDDDLLELEAEERPMEDDDOKAVKDELQDAGDEGL 2955
QY 958 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1017
Db 2956 DLEAEERPOLLEEDDDOKAVRDELQDGAADDDLELDETEERDLE-EGDDDEOKIKQKDEL 3014
QY 1018 DSDSDSGS---DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1046

```



```

Db 3015 HFDVDCALYEAQEKAAEADADAEDEED 3046

RESULT 14
US-09-134-001C-5080
; Sequence 5080, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5080
; LENGTH: 3696
; TYPE: PRF
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080

Query Match          4.9%; Score 279; DB 4; Length 3696;
Best Local Similarity 21.0%; Pred. No. 3.6e-07;
Matches 199; Conservative 134; Mismatches 351; Indels 262; Gaps 44;

QY 2 INKNNLLTKKKPIANK---SNRYAIRKFTVGTASIVIGATLLFGLGHNKAKEENSQ 57
Db 1264 IONNDATTEKEEVANLNVASQONVISKIDNATTNQIDIVSDGROSINATPDSIK 1323
QY 58 -----DVDSN-----TDELSDSNOSSEKENDVYINNQSINTDNNQII 99
Db 1324 RNAKNDIDIAADKKIKIQRINDATDEIDQAN-RKIEAKTEAKDNQFNSTDO--V 1379
QY 100 KKEETNNYDGIK-----RSEDRTESTVVDENEATFLOKTPODNTHLFEVEEKSS 152
Db 1380 NEATNGINKENTTPATYKSEAR-QAVONKANEQNIHIONTD--ATNEKEQEMIN 1434
QY 153 VESSNSSIDTAQPSHTTINRESVQTSVDNEDSHVDFANSKIKESNTESGKEENTEQ 212
Db 1435 RVSAELARVQO-----INAEHTTGQVKTIKDAITSL--SRINAVYKESARNAIEQ 1486
QY 213 -----PNKVEDS-----TTSQPSGYTNIDEXISQD-----ELL 242
Db 1487 KATQOTQFINNNDATDEKEEVANLVIATQKS-LDNINSLSSNNDENAKVAGINEIA 1545
QY 243 N-LPINEYENKARPLSTTSAPRSIKRVTVNOLAEOGSNVHLIKVTPDSITE-----G 295
Db 1546 NVLPATAVKSKAKKDIDOKLAQINOIOTHOTATTEEKEA--AIOLANKSNEARTATON 1603
QY 296 YDSEGVIAKADAENLIYDVFEDVDKYSQDITMTVDIK-----NTVPSDLTD-- 344
Db 1604 EHSNNGVAQA--KSGNHEIELVMPDAHKKSADAKO-SIDKNYNEQSNTINTP--DATDEE 1659
QY 345 -SFTIPRIKDSGELIATGTDNK-----NKQITFTFDVVKYENIKAHILKLSYIDK 397
Db 1660 KOKALDKIK-----IADAGYKNVDAQOTNOVSADANTEADITLTINQAVN----- 1705
QY 398 SKVNNNTKLDVEYKKTALSYVNT--ITVE-----YORPENNR--TANLQSMFTNIDKN 448
Db 1706 AKKSARVELDSKEDFKROINATPNATEEKKODAIQLNKRKDEKVL-----INQDRD 1761
QY 449 HTVEQ-----TIYINPLRYSAKETNVNISGNGEGSTIIDSTIIKIVYKGDNON 498
Db 1762 NEVDQHNKIGLQLEETIHANFTPKSDALOELQ-----TKFISQTELLINNNKQATNBE 1813
QY 499 LPDSNRIDYSEYEDYTVDDYAOQLGNNDVNNINFGNIDSPYIIKIKSYKPRNKDQYTIQ 558
Db 1814 KDEAKRLEISKNTTITNINOATNOVD-----NAKDNQMNIEIATII 1856

```

```

QY 559 QVTVMQTTINEYGEFRFASVNTIAFSTSGGCGDLPPEKTYKIGDYWEDYK---- 614
Db 1857 PATRIKT-----DAKTA-----IDKKAQ 1875
QY 615 -----DGIONTNDNEKPLSNVLTLYPDGTSKSVRTDEGKQFQDLK-NGLYTKITFE 668
Db 1876 QVTIINNNNDATDEKAEKARLYEKAKIE--AKSNITNSPDEREVNKAKTNGLE-KINNI 1932
QY 669 TPBEYTPTLKHSCTNPALDSGNSVWYTINGQDDMTIDSGYQPKSLG--NIVWYDTN 726
Db 1933 QPSTQYKT-----NAKQINDXAOEOLIQINNTPATTEERKEQEAATRVNAGIAQIIONIN 1988
QY 727 KDGIQDDEKISGVKTYLTKDENGNIISTTTDENGKYOFPDNL-----NSGNYVHFDPK 781
Db 1989 AHSQGEVNESKTNISI-ATIKSVQPNVIAKKPTAISLNLQEOANNOKTLGNGN----- 2039
QY 782 SGMQTQTTTDSGDDEQAD-----GEEVHYVTITTHDEFSDIN 818
Db 2040 -----ATDDEKEAKOLYTKLNEIOIKIHESTODNOVDN 2074

RESULT 15
US-09-134-001C-4054
; Sequence 4054, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4054
; LENGTH: 699
; TYPE: PRF
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4054

Query Match          4.9%; Score 278; DB 4; Length 699;
Best Local Similarity 21.1%; Pred. No. 4.8e-08;
Matches 152; Conservative 122; Mismatches 233; Indels 154; Gaps 32;

QY 15 IANKSNKYAIRKFTVGTASIVIGATLLFGLGHNKAKEENSVDYKDSNTDDELSDSNQ 74
Db 12 VKTRQNKSIKFSVGAASSILIALLPFGGSAQA-AEQQODKGVENSTTQSIGDEREK 70
QY 75 SSDEKNDYINNOSINTDD--NNQIIRKETNNYDGIK-----RSEDRTESTVVDENE 128
Db 71 LSEOSTQONKNVNEKSNVSDITENESLNETPKMEDLIQOQKDSQONKSSVVEQNEEN 130
QY 129 ATPLOKTPODNTHLFEVEEKSS-----VESSNSSIDTAQOPS--HTTINRESVQ 178
Db 131 EAFYKHSBEERPOEOVYLEKHAENNOTLHASKAQSHEVYKTRPSQDNTTAQOEDQK 190
QY 179 TSDNEDSHVDFANSKIKESNTESGKEENTIEOPNK-VKEDSTTSQPSGYTNIDEXISN 237
Db 191 ENLSKQDTQ--SKKTDLLRATQNGQKOSQSTEEVNEKVKND--TQOVTAKNDDQV-- 244
QY 228 QDELINLPINE-----YENKARPLSTTSAPRSIKR-----YTVNQLAAEQ 277
Db 245 --ETFNLSKEEPLKVDQANP--TTDKDKSSKNDKSHDGLANLESNAVAATTKQSKQ 300
QY 278 GSNVNLHILKVDOSITE--GYDSEGVIAKADAENLIYDVFEDVDKYSQDITMTV--DID 334
Db 301 VSEKNE--DQTNKSAKQOYKNNDEIILVHGFNGFTDIDINSVLTNHYGGKMKMIQDLE 358
QY 335 KNTVPSDLTDSFTIPIKIDNSGEI--IATGTVD--NKNKQITVYFTFDVVKYENI---- 385

```

```
Db 359 ENGVEYEAISAFSGSNYDRAVELYYIKGRVDYGAHAHAAGHERYKGYEGVYKDK 418
OY 386 ---KAHLKLTSTYIDKSKVPNNNTKLDVEYKTAHSVNTITVEYORPNENSTANLQSMFT 442
Db 419 PGOKIHL-----VGHSMGQOTIROLEELLRHGNPEVEYQ----- 453
OY 443 NIDTKNHTVEQTIYINPLKISAKE-----TNVNISGNGDEGTTIIDDSTIK--VYKVG 494
Db 454 ---KOHGE---ISPLFOGHDNMVSSITTLGTPHNGTHASDLGNEAIVROLAYDVG 505
OY 495 DNQNLDPDSNRIDYSEY--EDVTNDYAOLGNNNDVNINNGNIDSPYIIKISK----- 546
Db 506 KMYGNKDSRYDFGLEHMGKOKPNESTIO-----YKRVONSKLMSK 548
OY 547 ---YDPNRDDYTTIOQVYTMOTTI--NEYTGEFRTASYDNTIAFSTSSGOGDLPPEK 600
Db 549 DSGLDLTRDGATDLNRKTSLNPNIVYKTYTGE---STHKTLA---CKOKADLMFL 599
OY 601 TYKI-GDIYWEVDYK-----DGIQNTINDNEKPLSNVLVTLITPDGTSKSVRTDEDGKIOF 654
Db 600 PFTITGNLIGKAKEKEMRENDGLSVYISSQHPNOKYVEAT--DKNOKGVMYVTPTKHDW 657
OY 655 D 655
Db 658 D 658
```

Search completed: March 14, 2003, 13:01:18  
Job time : 45.6129 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 13:00:30 ; Search time 16.3082 Seconds  
(without alignments)  
3086.325 Million cell updates/sec

Title: US-09-147-405B-15  
Perfect score: 5646  
Sequence: 1 MINKKNLTKRKPIRANKSN.....FAGIGALLGKRRKRNKN 1092

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2-6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2-6/ptodata/2/pubpaa/PC7\_NEW\_PUB.pep:\*
- 3: /cgn2-6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2-6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2-6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2-6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2-6/ptodata/2/pubpaa/PC7S\_PUBCOMB.pep:\*
- 8: /cgn2-6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2-6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2-6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2-6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2-6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2-6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2-6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2069	36.6	1349	10 US-09-815-242-5898	Sequence 5898, Ap
2	2069	36.6	1349	10 US-09-815-242-13137	Sequence 13137, A
3	2011	35.6	932	10 US-09-815-242-5578	Sequence 5578, Ap
4	2011	35.6	932	10 US-09-815-242-12438	Sequence 12438, A
5	1634	28.9	1021	10 US-09-815-242-5471	Sequence 5471, Ap
6	1634	28.9	1021	10 US-09-815-242-12544	Sequence 12544, Ap
7	1441.5	25.5	841	10 US-09-815-242-5779	Sequence 5779, Ap
8	1441.5	25.5	841	10 US-09-815-242-12751	Sequence 12751, A
9	1023	18.1	265	12 US-10-073-256-78	Sequence 78, Appl
10	865	15.3	1113	10 US-09-815-242-5836	Sequence 5836, Ap
11	865	15.3	1113	10 US-09-815-242-12769	Sequence 12769, A
12	607	10.8	2344	10 US-09-815-242-12713	Sequence 12713, A
13	500	8.9	1001	10 US-09-815-242-12686	Sequence 12686, A
14	491.5	8.7	1018	10 US-09-815-242-5797	Sequence 5797, Ap
15	491.5	8.7	1018	10 US-09-815-242-12838	Sequence 12838, A
16	477	8.4	767	10 US-09-815-242-5899	Sequence 5899, Ap
17	477	8.4	767	10 US-09-815-242-13140	Sequence 13140, A
18	458	8.1	978	10 US-09-815-242-5456	Sequence 5456, Ap
19	453	8.0	560	10 US-09-815-242-13057	Sequence 13057, A

20	448	7.9	502	10 US-09-815-242-5904	Sequence 5904, Ap
21	382.5	6.8	2368	10 US-09-815-242-5635	Sequence 5635, Ap
22	382.5	6.8	2368	10 US-09-815-242-12389	Sequence 12389, A
23	382.5	6.8	2478	10 US-09-815-242-5816	Sequence 5816, Ap
24	382.5	6.8	2478	10 US-09-815-242-12967	Sequence 12967, A
25	349	6.2	345	10 US-09-813-820-7	Sequence 7, Appl1
26	322.5	5.7	1031	10 US-09-815-242-10932	Sequence 10932, A
27	320	5.7	665	9 US-09-820-843A-107	Sequence 107, App
28	295.5	5.2	102	10 US-09-864-761-44209	Sequence 44209, A
29	287	5.1	1400	10 US-09-764-176-7	Sequence 7, Appl1
30	286.5	5.1	2150	9 US-10-135-322-17	Sequence 17, Appl1
31	286	5.1	1060	10 US-09-955-909-9	Sequence 2, Appl1
32	277.5	4.9	870	10 US-09-815-242-5493	Sequence 5493, Ap
33	277.5	4.9	870	10 US-09-815-242-12637	Sequence 4, Appl1
34	271	4.8	1115	9 US-10-153-273-2	Sequence 2, Appl1
35	267.5	4.7	2122	9 US-09-813-214A-9	Sequence 9, Appl1
36	264.5	4.7	489	10 US-09-876-889-350	Sequence 350, App
37	263.5	4.7	1770	10 US-09-841-132-444	Sequence 444, App
38	260	4.6	1183	9 US-09-870-759-45	Sequence 45, Appl1
39	259.5	4.6	1435	9 US-10-153-273-4	Sequence 179, Appl
40	257	4.6	1776	10 US-09-841-132-179	Sequence 34, Appl
41	254.5	4.5	1139	10 US-09-801-574-34	Sequence 3, Appl1
42	252.5	4.5	1786	9 US-09-742-096-3	Sequence 1835, App
43	250.5	4.4	821	10 US-09-841-132-195	Sequence 5835, Ap
44	247.5	4.4	2434	10 US-09-815-242-5835	Sequence 12996, A
45	247.5	4.4	6281	10 US-09-815-242-12996	

## ALIGNMENTS

RESULT 1  
US-09-815-242-5898  
Sequence 5898, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes In  
FILE REFERENCE: ELITRA, 011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 5898  
LENGTH: 1349  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5898  
Query Match 36.6%; Score 2069; DB 10; Length 1349;  
Best Local Similarity 36.6%; Pred. No. 3.1e-92;

Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;

QY 1 MINKN-NLTKKPIANKSKYARKTMTVGTASIVATLFGIGHAEAEKSVODV 59  
Db 1 MLNENKATIRKGVNSRLKFSIRKTYGTASILVGTTLFGLGNEAKAEESTNEL 60  
QY 60 KDSMTDELSDNSOOSPEEKNDVJNNOSINTDNNOIKEEENNDGAEKSEDETE 119  
Db 61 NEAT-SSASDQSSD-KVDMQOLNEDNTKNDNO-KEM-----V3SQNETTS 105  
QY 120 STTVNENEAFLQKTPDNTHTLTHEEVKSSSVSSSSIDTAQOQPSHTTINRESVOT 179  
Db 106 NGNKLIERES--VOSTGNKVEVSRAKSDQASPEKSTEDINTKQ-----TISNEAQA 157  
QY 180 SDNVEDSVSPANSKIKESSTESKEKENTIEQPNKAKEDSTTSPSGTYT-DEKISNQ 238  
Db 158 PDLENKSVN-----VQPTNEENKVD-----AKESTT-----LVAKSPAIKSN 198  
QY 239 DELL-NLPINEYENKAPLSTSAQ-----SIRKVTNOLAEQSGSVNHLIKTQDSI 292  
Db 199 DELVDNNSNSNNENNAIDILPKSTAPKRLNRMRIAAVQSSSTEAKVNDLITSNTLT 258  
QY 293 TEGYDSEGVIAKHAENLIYDFEVDKVKSGDTMTVDIDKNTV-----PSDLTDSF 346  
Db 259 VVADKNNKIYPADQYLSLKSQIT--VDQVKSQGYFTIKY-SDTVOYGLNPEDIKN-- 313  
QY 347 TIPKTD-NSGEIITATGYDNKKNKOITFTFDYVDKYNIAHLKLTSTIDSKVPPNNT 405  
Db 314 -IGDKIDPNNGETIATAKHADANMLITFTDYDRFNSVQGVINYSIYMDATIP--VS 370  
QY 406 KLDVEKTAALSVNKTIVEYQRP--ENRATANLQSMFTNIDTKNHT-----VEQTI 455  
Db 371 KNDVEFNNTIGTTTKTANIOYPIYVNEKNSISAF--ETSHVGNKKNPGYKOTI 428  
QY 456 YINPLRYSAKETNVI-----SGNDEGSTIIDSTIIKVKYGNQNPDSNRIVDYS-- 509  
Db 429 YVNPENSLTNAKLKQVYHSSYPNNIGQINKDVIKIQYPKGYTL--NKGYDVATK 485  
QY 510 EYEDTNDYAO--LGNNDVNIQFNIDSPYIIKIVSKDPNKKDDVTTQJYTMQTT 566  
Db 486 ETTDTN-QYLOKITQYGNNSAVIDGNADSAVYVAVNKKFOYTNSPSTLWQMASTST 544  
QY 567 INEYGEFTASDYNTIAFSTSSGOGDLPPEKTYKIGDYWEMEDVDGIDQ----- 618  
Db 545 GN-----KSVSTGALGFTNNQSGAG--QEVYKIGNYWEMETNNNGVQLEKQVG 594  
QY 619 -----NTN----- 621  
Db 595 NVTYVFNNTMTKVGAVTEKEDGSYLLIPNLPGNDYRVEFNSMLPKGYEVPKQCNNEEL 654  
QY 622 -----DN-----EKPLSNVLT-- 634  
Db 655 DSNGLSVITVNGKDNLSADGIYKPKYNLCDYVWEDTNKNGIDQDEKIGISVTVTLKD 714  
QY 635 -----TYPDG----- 639  
Db 715 ENGDLKATVTTADADKRYFTDLHNGNYKEFTTPEGYPTTIVSGSDIEKDSNGLTTCV 774  
QY 640 -----TSKSV----- 644  
Db 775 INGADNMTLDSGFYKPKYNLGNVWEDTNKDKODSTFKGISGYTVTLKNGEVLQTT 834  
QY 645 KTDDEGKQFQDLKNGLYTKITFEPEGTPPLKSHSGTRPALDSEGN----- 692  
Db 835 KTDKDKQYQFTGLENG-YKVEFEFTPSGYTPQVSGSGTDEGIDSNGSTTGVIKDKNDT 893  
QY 693 -----VTI-----NGQ-- 700  
Db 894 IDSGFYKPYNLGDYVWEDTNKNGVQDKBEKISGYTVTLKDBNDKVLKTVTTDENGKIQ 953  
QY 701 -----DMTIDSQFYQTPK 714  
Db 954 FTDLNGFTYKVEFEFTPSGYTPSVTSGNDEKDSNGLTTCVTKADADNMTLDSGFYKTPK 1013

QY 715 YSLGNVWYDNTKNDIGDDEKIGSVKVTLLKDENGNITSTTTDENGKYQFDNLNSGN 774  
Db 1014 YSLGDYVWYDNTKNDIGDDEKIGSVKVTLLKDENGNITSTTTDENGKYQFDNLNSGN 1073  
QY 775 IVHFDPSGMYTQTTSDSGDDEADAGEVHTTTDHDPSIDNGYVDESDSDSDSDSD 834  
Db 1074 KVIFEKPRAGLQYVNTTDED-KDADGGEVDTTIDHDPLDKNYFEEDT----- 1123  
QY 835 SD 894  
Db 1124 --SD 1181  
QY 895 SD 954  
Db 1182 SD 1241  
QY 955 SD 1014  
Db 1242 SD 1300  
QY 1015 SD 1073  
Db 1301 -----TTKDHNRKAKALPETGSENGNSNATLFGLE 1332  
QY 1074 AGLALL-GRRRNRK 1089  
Db 1333 AALGSLLEGRRKQNK 1349  
  
RESULT 2  
US-09-815-242-13137  
Sequence 13137, Application US/09815242  
Patent No. US2002061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ECTTRA 011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13137  
LENGTH: 1349  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-13137  
  
Query Match 36.6%; Score 2069; DB 10; Length 1349;  
Best Local Similarity 36.6%; Pred. No. 3,1e-92;  
Matches 533; Conservative 168; Mismatches 280; Indels 476; Gaps 46;



QY	2	IKKKNNLLTKKKPLANKSNKAIKFKFVPASTAYIGATILFLFGJLHNAKAEBSYODYKD	61
Db	1	MNNKTAFTNRGNLIPNRINKFSIKRYSVAGTASILVGTLLIFJLSGHAKEA-----	52
QY	62	SNTDLELSDSDSDSDEEKNVNNNOSINMTDDNNQIIKKKEETNNVQGIKRSERDRETE	121
Db	53	-----HINEL-----NOSKN-----ET	65
QY	122	TNVDENEATFLOKTPPODNTHLTEEEVKESSVESSNSSIDTAQOPSHTTINREESVQSD	181
Db	66	TAPSENKTT-----EKVDSROO-----N	83
QY	182	NVEDOSHVDFAANSIKESNTFSSGEKENTIDOPKNVKRSDSTTSQPSGTINJDEKISNDEL	241
Db	84	NIEGSTISD--QPKVNEBDNTSVKE--TTEP-----ONTSTOPTKONN--DAMAKD--	131
QY	242	LNLPINEYENKARPLSTTSAPSK-----RVTVNOIAA--EOGSNNVHLIKYVDOGI---	292
Db	132	-----NLAAONISTQAKOVSTPKPTTIKPRTLRMAVNVTYAAQOGSTNNDKHFNINIDALD	190
QY	293	-----TEGYDDEBGVYKAKHAENLIYDTEFYDDKVKSGDVTWYDIDKNVPSDL	342
Db	191	KGHVNKTTGNTFPAATSDYLK-----LKANYTIDDSVKEGDTFTEFGYQFRRPSV	242
QY	343	TDSFTPIKONSEIATATGYDKNKNOITTEFDYDVKXENIKAHKLTSYIDRSKXVFN	402
Db	243	RLPSTONLILYAQONITLAKGIYDEISTTYYTTFNNYDQUTNINSFGEYVAFKRRNATP	302
QY	403	NNTKLDEYKTAALSSVVKTIIVEQRENERNTANLOSMFNTIDTKNHTVEQUTIYN--PLR	461
Db	303	DKTAPMEVITLGNKYSKANVIYD---GNKGQOOLISSTYININNEDLSRMITYVYVQPKR	359
QY	462	YSAKETNV--NISGNGDEGSTIIDSTIIKVKYGDNON-----LPDSNRKYDSEYEDV	514
Db	360	TYTKEETEVNTLT-----GYKFNPPAKNFKEIYEV--TNQNGVDSFTFDTSKLIDVTDKFKI	413
QY	515	TNDYDAAQGNNDVNNINFG--NIDSPYIIKVIKSYIPNRKDYTTIOQYTNQTTINERTG	572
Db	414	T---YSNDNKATAYDLNLGSSOSKQYIIIOOVAAPRMSSTDNCKIDYLTLETQNGKSMWN	470
QY	573	EERTASYDNITIAFTSSGOGGDLPPPKRYKIGDYVEDVKQGIONTNDEKPLENVLY	632
Db	471	-----SYENVAGSSSTAND-----QKKNLGDYVMEIDNRKAGKDA--NEKGIGYVY	516
QY	633	TLTYPDGTS--KSVRTDEGKYQDFGLNGLTYKITFEETPEGRYPTLKHSGTINPALDSEGN	691
Db	517	ILKDSNGKEIDRTTDTDENGKYOFTGLDNG--TYSVEFSTLAGYPTTANAGMTDAVDSGL	575
QY	692	SWMTINQODMTIDSGFYQTPRKISLGNLYWYPTNKGDIQGDDEKIGSYKYATLAKENGN	751
Db	576	TTTGVIKDAODMTIDSGFYKPKPKISLGDYWYYSNKNKGKSDSTEKGIDKQVYTLLENKE	635
QY	752	IISTTTDENKYOFDMLNNGSNYIVHEDKPSGMOQTDTBGDDDDADAGEEYHVTIHH	811
Db	636	VIGTTTKTIDENKRYFDMLDSGKYIVFEKAPAGLTQGTNTTEBD--KDAOBGEVDVYTIHH	694
QY	812	DDFSIDNGYDDESDDS	871
Db	695	DDFTLNGYEEH-----SDS	741
QY	872	DSDDS	931
Db	742	DSDDS	801
QY	932	DSDDS	991
Db	802	DSDDS	861
QY	992	DSDDSD	1051
Db	862	DSDDSD	892
QY	1052	K-LPDTGANEDYSGKTLGLGTFAGLGALL--GKRRKNRK	1089

```
Db      893 KALPETGSENNNSNGTTFGLFGAAGSLLEGRKKONK   932  
+ + + + : : | | | | | | | | | |  
  
RESULT 4  
US-09-815-242-12438  
Sequence 12438, Application US/09815242  
Patent No. US2002061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlson, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes In  
TITLE OR INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
PRIORITY FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/7191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12438  
LENGTH: 932  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12438
```

Query Match                35.6%, Score 2011; DB 10; Length 932;  
Best Local Similarity     42.9%; Pred. No. 1,2e+89;

	Matches	481:	Conservative	141:	Mismatches	278:	Indels	220:	Gaps	32:
OY	2	INKKNLLTKKKPKPIANKSKVAIKRFVTGTASIVIGATTLFGLGHNEAKAEENSVQDYKD	61	:	:	:	:	:	:	:
Dd	1	MNNKTATTNRKMIPNRRNKIFSRKYSVGTASILVGTTLFLSGHEAKAAE-----	52	:	:	:	:	:	:	:
OY	62	SNTDELSDSNQSSDEDEKNDVINNNOSINTDDNQIIEETINNYDGIKESSEDTEST	121	:	:	:	:	:	:	:
Dd	53	-HTNSEL-----NÖSKN-----ET	65	:	:	:	:	:	:	:
OY	122	TNVDENEAFTLOKTPQPDTHLTLEEVEKESSVESNSSIDTAQPSPSHITTINEESVOQSD	181	:	:	:	:	:	:	:
Dd	66	TAPSBNKTT-----EKVDHQO-----N	83	:	:	:	:	:	:	:
OY	182	NVEDSHVSDFANSKIKESNTESGKEENTTEOPNKVKEDSTTSOPSQGYNIIDEKISNOEL	241	:	:	:	:	:	:	:
Dd	84	NIEGSTSD--QPKNESDNDSVKE--TTEEP---QMTSVOPIRKÖNN--DAMANKD-	131	:	:	:	:	:	:	:
OY	242	LNLPTNEENKARPLSTSAPPSIR-----RYTVQLAA-EGGSVNHLIKYTDOSI---	292	:	:	:	:	:	:	:
Dd	132	-NLAAQNISTQAADVSTPTKTTYTIKPRLINRVAAVTVAAPAOGGTGVNKRKHFTNIDIAD	190	:	:	:	:	:	:	:
OY	293	-----TEGYDSEGVIKAKHAENLIIDVPREVDKKSQDTMVIDDKNTVPGDL	342	:	:	:	:	:	:	:
Dd	191	KGHVTKTGNTGFMAWTSDDLK-----LKANYTIDSQVKEGGTFFKYGÖFRPPQSV	242	:	:	:	:	:	:	:
OY	343	TDSFTIPRIKINSGEILITGTYYDNKNKOITTFPTYDVKYENIKAHMLITSYIDSKYPN	402	:	:	:	:	:	:	:

```
Db 243 RLSQOTNLNAGCNITAKIYDSESTTTTFTNNYDQITNIGSGEQAFARENATT 302
Qy 403 NNTKLDEYKATLALSVKNTITVEYQRPENENTANIQSMFTNIDRNHTEOITVIN-PLR 461
Db 303 DKRAYPMEVTLGNDRKSKNVIYDV---GNOKGQOLISSTYNINNEDLSRNNITVYVNOPKK 359
Qy 462 YSAKETNV-NISGNGDGSITIDSDTIKIKYKVDNON-----LPSNRITDYSEEDV 514
Db 360 TYKREFTVHJLT-----GYKFNPDANKFKIYEV-TNQNQFVDSFPTPKSLDVTDKFKI 413
Qy 515 TNDYQALGNNDVNIQF--NIDSPYIKISKYDPNKKDQYTIQQTVMQTTINEXTG 572
Db 414 T---YSNDKRTATVDLNGOSSSKQYIIQOVATPDNSVTNGKIDVTLEFQNGKSSNSN 470
Qy 573 EFRASYDNTIAFSTSSGQGGDPLPEKTYKIGDYWEDVDKDGIOQNTDNKEPLSNVLY 632
Db 471 -----SYSNNNGSSTANGD-----QKYNLGDYVWEDTKKKDKQDA--NEKIGKGYV 516
Qy 633 TLTYPDGTS-KSVRTDEDGKYQFDGLNGLTYKTFETPEGYPTLKHSGTNPALDSEGN 691
Db 517 ILKOSNGKELDRTTDENGKQYQFGTLGNG--TYSEFSTLAGYTPTTANAGTDADAVSDGL 575
Qy 692 SWMTTNGODMTIDSGFYQTPKSLGNYVWYPTNKGIOGDDEKGISGVVYTLKDENGN 751
Db 576 TTGCVIKADNMTLDSGFYKPKYSIGDYWYDKNKDGKQDSTEGKIKDYVYTLNKEGE 635
Qy 752 IISTTTDENGKYQFDNLSNGNYVHEFKPQMGOTTTDSGDDEQADAGEVHTYIDH 811
Db 636 VIGTKTKDENKRYFNDLSGKRYIFEKPAGLTQGTGNTEDD-KDADGGEVYVYTIIDH 694
Qy 812 DDESIDNGYIDDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 871
Db 695 DDFILDNGYEEET-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 741
Qy 872 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 931
Db 742 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 801
Qy 932 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 991
Db 802 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861
Qy 992 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1051
Db 862 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 892
Qy 1052 K-LPDTCANEDYGSKGTILGTLFAGLGLALL-GRRKRNK 1086
Db 893 KALPETGSENNNSNNGTLFGGLFALGSLILFGRKKONK 932

RESULT 5
US-09-815-242-5471
; Sequence 5471, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsten, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3471
; LENGTH: 1021
; TYPE: PR
; ORGANISM: Staphylococcus aureus
US-09-815-242-5471

Query Match      28.98; Score 1634; DB 10; Length 1021;
Best Local Similarity 37.48; Pred. No. 1.8e-71;
Matches 436; Conservative 155; Mismatches 334; Indels 240; Gaps 37;

Qy 18 KSNKYAIRKFTVGTASTIYIGATLFG--LGHNEAKAEKSVQDVADSNITDDELSDSNQSS 76
Db 5 KKEKHAIRKKSIGVAVSLVGLIGFGLLSKREADASENSV-----TQSDSASN 52
Qy 77 DEEKNDVINNQSINTDNNQIIEKEETNNYDGLERESEDRTESTNVNDEKATFLQKP 136
Db 53 ESKSSDSSVAAAKRTDNTNVDKRTTSNTNG-----ETSAQNPQ--QETT 99
Qy 137 QD-NTHLTHEEVK---ESSVSESSNSIDTAQOQSHHTINREESVQTSNDVNHVSDEFA 192
Db 100 QASNTNATEPTVGTGATATATKQAMPATQSSNT--NAEELVQTSNETAS----- 151
Qy 193 NSKIKESTEGKEENTIEQPNKYKEDSTQPSGYTNIDK--ISNDELNLPINEYE 250
Db 152 NDTNVTSSVNSPQNSTNAENVSTQDSTETATPSNSESAPSTDSASND-VVNOAVN--- 207
Qy 251 NKARPLSTSAQPSIKRTVNLQAAE--QGSNVHLIKYVDQSTIEGDESEGVYKAD 307
Db 208 -----TSA-PKRAFSLAVAADAPAGTDITN--QLTDVAVT--IDSGTIVYRHQ 253
Qy 308 AENLYDVTFEVDKVKSGDITVYDIDKNVPSDLTDSFTLPKIKDNGSGEIIAFGYDNK 367
Db 254 AGYVKLNGYFVPSNVAAGVFETKIVPKELNMGVSTAKYFPIWGD-QYLANGVIDS- 311
Qy 368 NKQITTYTVDVYDKENKAKHLKLTSTYDKSKVPNNKTL-DVEYKTL--SSVNTITV 424
Db 312 DGNVLYTFTDVTREKENTANITMPAYID---PENVTKGNVYLTGIGTNTASKYVLI 367
Qy 425 EYQRPENENTANLQSMFTNIDRNHTEOITVINLFRSAKETNV-----NISNGDE 477
Db 368 DYKYGQFHNLSIKGTIDQIDKTNNTYROTIVVN-----SGDNVYLPALGNLIPNKS 422
Qy 478 GSTIIDSTIIKYYKVGQDNQMLPDSNRITYDYSEEDVYNDYQALGNNDVNIQF---G 533
Db 423 NALIDAKNTDIAKYRV-DNANDLSGYYVNPSPDFEDVYNQVRISPPNANQKVEHPPTDD 481
Qy 534 NIDSPYIKISKYDPNKKDQYTIQQTVMQTTINEYGER--RTASTDNTIAFSTSSGQ 591
Db 482 QITPPYIVVNGHIDPASTG-----DLALRSTFGYSDNIFMSMSMDNEVAFNGSGS 535
Qy 592 GQG-DLP--PEKTYKIGY--VWEDVDKDGIONTNDNKPPLSNVLYTLTYPDGSKSVRT 646
Db 536 GDGIDKPVVPOPEGIEPIPEDSDS-----FGSDGSDSN 574
Qy 647 DEDGKYQFDGLNGLTYKTFETPEGYPTLKHSGTNPALDSEGNVWYTTINGOD-DMTI 705
Db 575 SDSC-----SDSGSDSTSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 607
Qy 706 DSGFYQTPKSLGNYVWYPTNKGIOGDDEKGISGVVYTLKDENGNIISTTTDENGKYQ 765
```

[illegible]

Matches	456:	Conservative	155:	Mismatches	334:	Indels	240:	Gaps	377:
QY	18	KSNNKAIKRFVTGASIVAGIATLFG-IGHNEKAKEENSVDYDKDSNTDDELSDSNDSS	76						
Db	5	KEKHAIKRSIGVASIVGLIGLGLSSKREADASENSV-----TQSDASN	52						
QY	77	DEEKDVINNNOSTITDQNNQIILKEELNNIDGIEKRSBDRTESTTNDENPATLQKPT	136						
Db	53	ESKSSDSSSVNAPKPTDPTNVSQDRTTSTNTNG-----ETSAQNAQ--QETT	99						
QY	137	QD-NHLEEEK-----ESSVSESSSSIDTAAQPSHTTINREESVQTSQNVDSHVSDPA	192						
Db	100	QASATNATEETLPVGEATTATKQANTPATITQSSNT--MAELVNQTSNEFAS-----	151						
QY	193	NSKIESNTESGKEENTIEOPKPKVEDSTTSPQSGYTNDK--ISNDELINLPINEYE	250						
Db	132	NDTNVSSVNSQNSQNTNMEVNSTQDTEATPKNNSESAPOSTDASQNMND-VYNQAVN---	207						
QY	251	NKARPLSTTSAOPSIKRYTVNOIAAE--QGSNNVHLIKYVDQSTIEGYDSEGYIKAH	307						
Db	208	-----TSA-PRMRAESLAAVADAAPAGDITN--QLTDVKT--IDSGTIVYPHQ	253						
QY	308	AENLYDYTFEVDKVKSGDPTTVIDDKNTVPSQDLDTSTIKIKINDSGEIIATGTQDK	367						
Db	254	AGYVLKNGFSPNSAVKVDFTKIVPEKELNNGVTSTAKVPPIWGD-QVLANGLID-311							
QY	368	NKQITFTFTDYDVKAEINIKAHKLTYSIDKSKVPPNNKTL-DVEYKTLA--SSVKTTIV	424						
Db	312	DGNVITFTFDYDKENYATNITMAIYD---PENVTKTGVTLTGTGCTFAKTVLI	367						
QY	425	EYQRPENENTANLQSMFTNIDTKNHTVEQTIYINPLRSAKETNV-----NISGNDE	477						
Db	368	DYEKGQGFHNLISIKTIDQIDKTNNTYRQTIYVNP-----SGDNVLPALTLGNLIPNKS	422						
QY	478	GSTIIDSTIIKVVYGVQDQNLPSDRIRYDEEYEDTDDOVAQLGNNDVAINF-----G	533						
Db	423	NALIDAKNTDIKVIYR-DNANDLSSEYIVPSPFEDVYVQVRISEPPNAQYVEEPTDD	481						
QY	534	NIDSPYIIKIVISKYBPKNKDYTTIQITVMQTTINEYGEF--RTASYNNTIAFSYSGQ	591						
Db	482	QITTYIYVNVNHHIDPASG-----DLALRFTFYGDENFTMRKSMNMEVAFNNGSS	535						
QY	592	GCG-DLP--PEKTYIKGDY--VWEDVDKDGIGONTNDNEKPLSNVLVTLTYPDGTSKSVRT	646						
Db	536	GGGIDKPVYPEQPDPEGEIEPIPEDSDS-----PESDSCSDSN	574						
QY	647	DEDGKYQEDGLKNGLTIKTTFETPEGYTTLKHSQTNPLADSEGNSVWYTINGQD-DMTI	705						
Db	575	SDSG-----SDSGSDSTSDSDSSASDSDASDSDSAS	607						
QY	706	DSGFTQPKYSLGNYVYDNTKDGIOGDEKISGVKVTILKQENGIIISTTTDENGKQ	765						
Db	608	DS-----DSTSDSDASDSDASD-----SDSASDSDASD	638						
QY	766	FDNLNSGNVIYHFDKPSGNTQTTTDSCGDEDEQADAGEEYHVTITDHDPSIINGVYDES	825						
Db	639	SDSASDS	697						
QY	826	DS	885						
Db	698	DS	757						
QY	886	DS	945						
Db	758	DS	817						
QY	946	DS	1005						
Db	818	DS	877						
QY	1006	DSDSDSVSD	1046						
Db	878	DSDSDASD	937						







```

RESULT 10
US-09-815-242-5836
; Sequence 5836, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5836
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5836

Query Match      15.3%; Score 865; DB 10; Length 1113;
Best Local Similarity 26.2%; Pred. No. 1,9e-34;
Matches 344; Conservative 174; Mismatches 365; Indels 430; Gaps 56;

QY 6 NLLTTRKKPIANKSNKAIKRFVTGASIVIGATLLFGLGHNEKAKEENSVDVDSYTD 65
DB 2 NKVTEAPTAETNK-----ATETPA--VEDTNAK 32
QY 66 DELSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 101
DB 33 ---SNSNAQPEETERTQVDTAKDLYKKSEVTEAKAEIKYLPKDISNLSNEIKKIA 89
QY 102 -----EENNVNDGIEKSEDTES-----TTVNDENAEFLQKTPQDNT----- 140
DB 90 LSEVLKETAENKEMQAPRTFRSVSSNARTVNVNS-ATAALAAADDTYTKGCTGNFAG 148
QY 141 HLTREEVKSSSVSSNSIDTAQOPSHTTINREESVOTSDVEDSHVSDF-----ANS 194
DB 149 DIIRKTYKEEPNECKLTAFNTNRPDTGT---KGALEYNDKIDN--KDFITTYPVANN 203
QY 195 KIKESNT-----ESGKEEMTIEQPNKYKEDSTTSPSG-----YTNIDKIS-- 236
DB 204 --NGCNTGADGWMFMFTQNGODPLNG--GILRDKGMANAGFKIDTAVNANNKVDKL 260
QY 237 NQDELNL-----PINE----- 248
DB 261 DADTNNLSQIGAAGVGVGTFVKGAGDVTNOVGONALNTRKDPVKNKIITYADNTNHLDG 320
QY 249 -----YENKARPLSTSAQPSIKRVTVNOLAEOGSMVNLIKVTDOSITEG 295
DB 321 QFHGQRLNDVYLNDAATSTITATYAGKTWK-ATTDDGIDXSQKYNFLI--TSSHMONR 377
QY 296 YDDS-----EGVIKHADEANLIYDVF-----EYDDKVKSGDTMTV-----DI 333

```

```

DB 378 YSNIGMRNLEGVITTPQADLIDVETVKQPIPKHKTIREPFTLEGSPDVIYQKEDG 437
QY 334 DKNVPSDLNDSFPIPKDNGSELIATG---TYDNKN-----KQITTFYTYVD 380
DB 438 EKTIT-----TPTKVDPPTGDVVERGEPTTEVTNKNPDEIYHFAPEEYPOGHKBEFD 489
QY 381 K-----YENIKAHKLKTSYID-----KSKVYNNNTKLDVE 410
DB 490 PNLPIDGTEEVPGKPGIKNPETGKVTMPVDDVTKHGPKAGEPEVTKKEIIPFEKKR---E 546
QY 411 YKTALSSVKNKITYEYQRPENRNTANLQSMFTNIDTKNHYEQITTYINPLRYSKERNVN 470
DB 547 FNPDLKPGEEKVQTQEQ-----TEKTTPTPTTINPL-----TGER 582
QY 471 ISGNGDEGSTIIDSTIIKVKYKVDNONLPDSNRHYDSESE-----DVTNDVYAOGLNN 525
DB 583 V-GEPTTEVTKPEPVDLITFG-GEVPOGKH---DEFDPMLPIDGTEEVPGKPGIK 635
QY 526 N-----DVNINFGNIDSPYIIVISKYD-----PNKDYT---TIOQ 559
DB 636 NPETGKVYTPPVDDVTKHGPKAGEPEVTKKEIIPFEKKRPNPDLKPGEEKVTOGSGTGEX 695
QY 560 TVTMQTTINETYGERPTASIDNTIAFTSSGQGGDLPEPEKTYIGDYWEDVDKQIGION 619
DB 696 TTTPTPTTINPLTGE-----KVGEPEPTTEVTKPEVDEITFGGGEVPO-CHKD 742
QY 620 TNDNEKPLSNVLTLYPDGTSKSVRTDEGKYQDFGDKNGLTYKTFEPTEGTYPTLKH 679
DB 743 EFDNRLP-----DGTG-----EVPGK---PGIKNETGKVYTPPVDDVT---KH 781
QY 680 SGTNPALDSEGNVWVTINGODMTIDSGFYQTPKYSLGNYVWYDTNKGIOGDEKGIS 739
DB 782 GPK-----AGEPEVTKKEIIPYETKR-----VLDPTEMPG-----S 811
QY 740 GVKYTLKDNENIITTTDENGKYQFONLNSGNTIVHFKPSCGTFOTTTSGDDEDA 799
DB 812 PDKVAQKGENEKTTPPTTIN-PLTGERVGE-----EPT--TEVTKPEIETIYVYA 861
QY 800 DGEVHYVITTDHDFSIDNGYDDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 859
DB 862 PEIIPHGTREE-----IDPNLPPEGETKYIIPGKDLKDETEIIEEPDEVIYIHGAKDS 916
QY 860 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 919
DB 917 DSDADSDSDADSDSDSDSDSDSDSDSDSDSDSDADSDSDADSDSDADSDSDSDSDSD 976
QY 920 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 979
DB 977 DSDADSDSDADSDSDSDSDSDSDSDSDSDSDSDADSDSDADSDSDADSDSDSDSDSD 1036
QY 980 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1039
DB 1037 DSDADSDSDADSDSDSDSDSDSDSDSDSDSDSDADSDSDADSDSDADSDSDSDSDSD 1063
QY 1040 DLGNSSDKSTKDKLPDTGANEDYGSKTLLGLFAGLAL-LGKRRKRRK 1091
DB 1064 ---DKTDKPNKKELPDTG--NDAQNNGTLFGSLIRALAGLFLVGRFRKKNKNE 1111

RESULT 11
US-09-815-242-12769
; Sequence 12769, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard

```

```

OY 147 VKESSSVSSSSSIDTAOOPSHTTINREESVOTSNVEDSHVDFANSKI-----196
D 17 LRDGMAVASOFKIDITAVN-----NVANGKDXDLADKATNNLSQGAAKVYGCFEVKNKA 70
OY 197 -KESNTBEGKKEENTEP-AK-VKEDSTSOPSGTINIDEKISNODELINPEINEYENKA 253
D 71 DGVINOVGOMALNKKDKKPVNKKIYADNTNNHLDGOFH-----GQRLLNDVLYLVNDAT 122
OY 254 RPLSTSAOPSIKRYTVNQALAGEGSNVNHILKVDOSITEGVDDS-----EGYKAMD 307
D 123 STIATVAGKRWK-ATTDDLGDKSQKNFL-TSHMOKNYSNGIMRTNLEGTITTP 179
OY 308 AENLIVDTF-----EVDDKVASGDMV-----DIDKATVPSDLTDSFTTIPKI 351
D 180 QADIDIDVEVTKOPIPHKTIREFPDTEPEGSDVIVOGGEDEKITT-----TPTKV 231
OY 352 KDNSEIATG--TYOKN-----KOITTFIDYVK-----YE 383
D 232 DPTGDVYEREPEPTTEVTKNPVDELVHAPEVPOGKHDEDPNLPIDIGTEVPCKPGIK 291
OY 384 NIKAHKLTSYID-----KSKVPNNRTKIDVEYKATLSSVNTTIVEYOR 428
D 292 NPETGKVVTPEVDDTKHGPKAGEBEVKEKEIPEKKR--EFNDLKGEEKVMOEGQ- 347
OY 429 PNEKRTALQSMFTNIDPKNHTVEGTIYNPLRYSAKETNVNISNGDEGSIIIDDSTII 488
D 348 -----GKETTTPPTINPL-----TGEKV-GEG-EPTEVKEPEVD 382
OY 489 KVEYVDNQNPDSNRIDYSEY-----DVYNDDYALGNNN-----DVNIN 531
D 383 EITQFG-GEVPOGKH-----DEFDNPLLIDIGEELVPGKGINFETGKVVTPEVDDVYKH 437
OY 532 FGNIDSPIIIIVISKYD-----PNKDDYT---TIQOTVMOTTINETYGEFRTA 577
D 438 GPKAGEPVTEEIEIRFEKKREPNPLDKGEEVTOGEGTETTTTPPTINFLTGE--- 493
OY 578 SYDNTAISTSSGOGGLPREKTYKIGDYVEDVDKOIGONTNNEKPLSVLVTLTP 637
D 494 -----KVEGEPTTEKEPVDEITOFGGEELVPO-CHKREFPNPI----- 534
OY 638 DGTSKSVTDEDKYOPDLKNGLTYKITTEETPEGYPTLKGSTNPALDSGNSVWTI 697

```

```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12713
LENGTH: 2344
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12713
Query Match 10.8%; Score 607; DB 10; Length 2344;
Best Local Similarity 24.1%; Pted. No. 1,1e-21;

```





```

      PRIOR APPLICATION NUMBER: 60/206,848
      PRIOR FILING DATE: 2000-05-23
      PRIOR APPLICATION NUMBER: 60/207,727
      PRIOR FILING DATE: 2000-05-26
      PRIOR APPLICATION NUMBER: 60/242,578
      PRIOR FILING DATE: 2000-10-23
      PRIOR APPLICATION NUMBER: 60/253,625
      PRIOR FILING DATE: 2000-11-27
      PRIOR APPLICATION NUMBER: 60/257,931
      PRIOR FILING DATE: 2000-12-22
      PRIOR APPLICATION NUMBER: 60/269,308
      PRIOR FILING DATE: 2001-02-16
      NUMBER OF SEQ ID NOS: 14110
      SOFTWARE: FastSeq for Windows Version 4.0
      SEQ ID NO 12838
      LENGTH: 1018
      TYPE: PRT
      ORGANISM: Staphylococcus aureus
US-09-815-242-12838

      Query Match      8.7%, Score 491.5, DB 10, Length 1018;
      Best Local Similarity 19.5%, Pred. No. 1.5e-16;
      Matches 232; Conservative 191; Mismatches 490; Indels 275; Gaps 141;

      5 KNNLLTKKKPKIANKSKAIAKRFYGTGTAISYVAGALLFLGLGN-EAKAEENSVDQKDSN 63
      2 KNNL-----RGIKRKHKGAAVSFLGTVITVVGQDGKPAALASEQKTTTVEENG 49
      64 TDDELSDNDGSDDEKNDVINNNOSINTDDNNQIIRKEETNNYDGIRKSEDPTESTTN 123
      50 -----NSATDNKTSETQTATTAVNHIEEQSYNAVTEQPSN-----ATQ 89
      124 VDENEATEFLQKTPDQNTHLTEEVKESSSESSNSIDTAQPSHTTINREDSVOTSDNY 183
      90 VTTEAPRAVAQAPCAQAPANIEVTEKEEVYKEEAKPOY-----KETTSODNS 136
      184 EDSHVSDFRANKSIRKSNSTESGKEENTIEDPNKVKEDSTTSQPSGYTNIDKISNDELIN 243
      137 GQROVDLTPKATNOVAETQE--VAQPRASE----- 169
      244 LPINEYENKARPLSTSAQPSIKRYTVNOLAEOGSNHNLIKTVDSITGEGYDSEGYI 303
      170 -----SKRPVRSADVALEAKESN-AKYETGTVYSKYTVELGI-EDHNNTNKV- 217
      304 KAHDAENLIYDVFVEDKVKSGDITVYDIDKNVPSDLTDSFTTPKIKDS-----GEI 358
      218 EPHAGQRAVLKYLKFEENGLHOGDYFDEFTLSNNVTHGVSARKVPEIKNGSVYMATGEV 277
      359 IATGYIDKNKNOITFTFTDYVDKYENIKAKHLITVYIDSKYPRNNNTKLDVEYKALSSV 418
      278 LEGG-----KIRFTFTNIDEDKVDVTALEINLEIFDEKTYVOTNGH-----QTTSFL 324
      419 N-----KTIIVEQOREPNEKNTANLQSMFANIDTKNHTVQOTIYINPLRYSAKETVNIISG 473
      325 NEEQTSKEIDVYKKQSIGYVYANLNSITFPKANRRESHVAFIKR--NNGGTTISTVYIG 382
      474 NDEGSTIIDSTIIKUYA-VGDNONLPDSNRIY--DYSEEDVYNDVYQOLG--NNN 526
      383 TLMKGSNONGNQPKVAFIEYFLGNEDIAKS--VYANTDTQSKFEVTSNMNGNLNONG 440
      527 DVNINFGNIDSPYIIVKISKYDPNKD--DYTTIQQVYVMTQITINEVTGEFRASVYNTJA 584
      441 SYSLNLEINDKTYVVHYHDEYELNGTDEVDFRT-QVNGHPEQLYKYYDDRGVYTLTWNGVL 499
      585 FGRSSQGGQGDLPPEKTYTIGDVVMEVDVQDG--IONTNDNEKPLSNVLVTLLYPPGTSK 642
      500 LYSNKRANGK-----KNGPITIONNK-----FEYKEDTJK 528
      643 SVRTDEGKVFQEDGLKNGLYTKITFPETPGQYPTTLKHSNTPALDEGNSV--WVTJNG 699
      529 EFTL---GGYD---KNLVT---TVE--EEDYSTLIDYHRAIDGCGGVYDGYETIEE 576
      700 QDDMTIDSGFYQTPKISLGNVYVYDINKKQIQGDDDEKGISGVKVTILKDNENGINISTTTD 759

```

[illegible]

Search completed: March 14, 2003, 13:08:12  
Job time : 31.3082 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01 ; Search time 22.1792 Seconds  
(without alignments)

4733.207 Million cell updates/sec

Title: US-09-147-405B-15

Perfect score: 5646

Sequence: 1 MINKNNLLTKKKPIANKSN.....FAGLGALLGRRRRKKNNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5646	100.0	1092	T30214	fibrinogen-binding
2	2698.5	47.8	1166	T28680	fibrinogen-binding
3	2589	45.9	1141	E89624	hypothetical prote
4	2234	39.6	1385	D89824	hypothetical prote
5	2093.5	37.1	953	C89824	hypothetical prote
6	1880	33.3	1315	T28679	fibrinogen-binding
7	1652	29.3	989	D89852	fibrinogen-binding
8	1628	28.8	933	S41539	fibrinogen-binding
9	1536	27.2	877	F90070	Clumping factor B
10	785	13.9	882	AG1671	probable peptidogl
11	782.5	13.9	903	AG1299	hypothetical prote
12	624.5	11.1	2271	F90073	hypothetical prote
13	554	9.8	406	S38170	Srpf40 protein - ye
14	549	9.8	3394	T18501	hypothetical prote
15	549	9.7	2570	T17451	fimbriae-associate
16	525.5	9.3	1038	H90053	hypothetical prote
17	520	9.2	940	S19702	fibrinectin-bindin
18	502.5	8.9	961	G90053	hypothetical prote
19	487.5	8.6	1018	A32192	fibrinectin-bindin
20	471	8.3	334	A54138	acidic repetitive
21	471	8.3	640	A54502	S antigen precursor
22	452	8.0	2910	T28156	DNA-directed RNA p
23	442.5	7.8	1999	AB2018	hypothetical prote
24	435	7.7	1192	A71623	probable secreted
25	430.5	7.6	4550	T18440	hypothetical prote
26	391.5	6.9	1217	S52714	sericinib - silkwo
27	388.5	6.9	695	S27390	calcium-binding pr
28	381.5	6.8	955	T18435	hypothetical prote
29	379.5	6.7	3724	T18427	hypothetical prote

30	371.5	6.6	1305	2	T00670	der--
31	370	6.6	1063	2	D86731	hypotheical p...
32	368	6.5	1664	2	T18262	lustrin A - Califo
33	366	6.5	3844	2	T18402	fmbb protein (limp
34	364.5	6.5	792	2	T42963	hypothetical prote
35	363.5	6.4	1360	2	T18403	hypothetical prote
36	362	6.4	1193	2	G71605	cell wall surface
37	351.5	6.2	1428	2	T08852	serine/threonine-s
38	345.5	6.1	2481	2	D90011	probable membrane
39	343.5	6.1	1658	2	S55101	S64942
40	338	6.0	4776	2	E95206	hypothetical prote
41	336	6.0	2485	1	H71621	hypothetical prote
42	335.5	5.9	4910	2	A36811	aglutinin-like ad
43	333.5	5.9	1072	2	A86827	
44	333	5.9	797	2	A36811	
45	331.5	5.9	1419	2	T30531	

#### ALIGNMENTS

##### RESULT 1

T30214 fibrinogen-binding protein - Staphylococcus epidermidis

C:Species: Staphylococcus epidermidis

C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000

C:Accession: T30214

R:Nilsson, M.; Frykberg, L.; Floock, J.L.; Pel, L.; Lindberg, M.; Guss, B.

Infect. Immun. 66, 2666-2673, 1998

A:Title: A fibrinogen-binding protein of Staphylococcus epidermidis.

A:Reference number: Z20781; MID:98261511; PMID:9596732

A:Accession: T30214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11092 <NID>

A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAI76638.1

Query Match 100.0%; Score 5646; DB 2; Length 1092;

Best Local Similarity 100.0%; Pred. No. 2.9e-211;

Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MINKNNLLTKKKPIANKSNKYAIRKFTVGASIVIGATLLFGIGHNKAKEENSVDYK	60
DB	1	MINKNNLLTKKKPIANKSNKYAIRKFTVGASIVIGATLLFGIGHNKAKEENSVDYK	60
QY	61	DSNTDDELSDSDSDSEKNDVYNNOSINTDNNQIKKEETNNYDGIRKSEDPRES	120
DB	61	DSNTDDELSDSDSDSEKNDVYNNOSINTDNNQIKKEETNNYDGIRKSEDPRES	120
QY	121	TTNDENATFLQKTPQDNTLTFEEVKESSSVSSNSIDTAQPSHTTINRESYOTS	180
DB	121	TTNDENATFLQKTPQDNTLTFEEVKESSSVSSNSIDTAQPSHTTINRESYOTS	180
QY	181	DNVEDSHVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGTYNIDKISNO	240
DB	181	DNVEDSHVSDPANSKIKESNTSGKEENTIEQPNKVKEDSTTSQPSGTYNIDKISNO	240
QY	241	LNLPINIEYENKARLSTTSQAQPSIKRTVNOALAEQSSNNHLIKYDOSSTEDYDSE	300
DB	241	LNLPINIEYENKARLSTTSQAQPSIKRTVNOALAEQSSNNHLIKYDOSSTEDYDSE	300
QY	301	GVTKAHDAENLIYDVTFEVDKVKSGDTMTVDIDNTVPSDLTDSFTIPKIKDNGEITA	360
DB	301	GVTKAHDAENLIYDVTFEVDKVKSGDTMTVDIDNTVPSDLTDSFTIPKIKDNGEITA	360
QY	361	TGTVDNKKKQITFTFDYVDKYEIKAHKLKTSYIDKSKVPNNNTKLDEYKTAALSVNK	420
DB	361	TGTVDNKKKQITFTFDYVDKYEIKAHKLKTSYIDKSKVPNNNTKLDEYKTAALSVNK	420
QY	421	TTTVEYGRPNERNRANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETVNNVSGNDESGT	480
DB	421	TTTVEYGRPNERNRANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETVNNVSGNDESGT	480





```

Db      61  NATT-----SASDNOSD--KYDMOOLNQBEDNTKNDQ-----KEVSSQONE 102
Qy      120 STTNVDN-EAFLQKTPQDNHILTEEVKSSSVSSSSSIDTAKQPSHTINPEEQ 178
Db      103 TTSNGKSTLEKSSVOSTGKNGKVESTAKSDEQASPKSTVEDNTKQ-----TISNOEGIQ 157
Qy      179 TSDNVEDSHVDFANSKIKESNTESGKEENTIEOPN-KKEDSTTSQPSGYINDEKISN 237
Db      158 -POLLENKSVN-----VQPTNEKKNVAKTESITLANKSAIKS-----NAELVYN 205
Qy      238 QDELLNLPINEYENKARPLISTSAOPSTIKRVYNOLAAEQS-----NNHLIKVTDOST 293
Db      206 NSNSNN-----ENNADILPKSTAPKSLNTRMAAIOQPSNDSKNNVDLITSMITLIV 259
Qy      294 EGYDSEGVIAKHADENLIDVTEFVDOKVSGDTMVDIDKNV-----PSDLTSDT 347
Db      260 VDADSKSTIVPQDYLKSLQIT--VDKVKSGDYFTIKY--SDTYVYVGLNEDBEIKN--- 313
Qy      348 IPKIRD-NSGEIATGYDNKNKQITTEFTDYVDKYEKNIKAKHLKLSYIDKSKVFNNTK 406
Db      314 IGDIDPNNGEITATAKHTANNLITFTDYVDKFNKYMGINSTYADDTIYVD-K 371
Qy      407 LDVEYKTAALSSVNTIYEQRP--NENFTANLQSMFTNIDTKNH-----VEQRTY 456
Db      372 KDVPESVITIGNQITTTADITPAYKADNNSIGSAFT--ETVSHVGNVEDPGYINQYVY 429
Qy      457 INPLKSAKETVINISGNDEST-----IIDSTIIKYKVDONLDPDSNRIYV--SE 510
Db      430 VPMKDKLGAUKLKEAYHPKPTINIGQINQVNTIKIRVEGYTL--NKGQVNTND 486
Qy      511 YEDVYND--DYAQLGNNDVNINFGNIDSPYIKIYKSKYDPKDDYTTIQOQVNTMOTIN 568
Db      487 LYDVIDEFPKNKNTYGSNOSVNLDPCDITSAIYVNMNTKFOYINSSPILYQMATLSSIGN 546
Qy      569 EYTGEPFRASTIDNTAESTSSGQGGDLPPEKTYKIGDYWEDVEKDGIO----- 618
Db      547 -----KSVSTGNALGFTNNOGAG--QEVYKIGNVWEDTKNGQVDELGEKGVGNV 596
Qy      619 -----NM----- 621
Db      597 TYTVEDNNNTKVGAVTKEDGSYLIPLPLNGDYVEFSNLPKGYEYTPSKOGNEELDS 656
Qy      622 -----DN-----EKPLSNVLTL----- 634
Db      657 NGLSSVITVNGKDNLSADLGIKPKRYNLGDYWEDTKNGIDQDEKIGSGVTLKDN 716
Qy      635 -----TYPDG----- 639
Db      717 GNVLTVTVDADGKYKFTDLNNGNYKVEFTTPEGYPTTVTSGSDIEKDSNGLTTGVIN 776
Qy      640 -----TSKSY-----RT 646
Db      777 GADNNTLDSGFKTKPKYKLYNLGYWEDTKKDGKODSTEGKISGVTYTLKNGEVLQTTKT 836
Qy      647 DEDGKYQDGLKNGLTLYKITEETPEGYPTTLKHSSTNPDSEGS----- 692
Db      837 DKDKGYQTGLENG--TYKVEFETPSGYTPTQVGSSTDEGIDBSNGSTTGVIKDKNDITD 895
Qy      693 -----VW-----VTI-----NGO----- 700
Db      896 SGFYKPYNLGDYWEDTKNGKNGVODKDEKIGSGVTLKDNMDKYLKTVTTDENKRYFT 955
Qy      701 -----DQMTIDSGFYQTPKYS 716
Db      956 DLNNGTYVEETPSGYPTVTSQNDTEKDSNGLTTGVKADANMTLDSGFKTPKYS 1015
Qy      717 LGNYVYDTPNKDGIQDEDEKIGSGVTLKDNNGNIISTTTTDENGKYOPDLNNGNIV 776
Db      1016 LGDYWYMSNKGDKDSTEGKIKDYKVLNLEKGEVIGTITTDENKGRFNDLDSGKRYV 1075
Qy      777 HFDPKSGMTQTTTSGDDDEQADAEVHVTITDHDPSIDNGYVDE--SDSDSDSDSDS 835

```

```

Db      1076 IFEKPTGLTGTCTNTTEDD-KDADGGEVDVYITIDHDDFTLDNGYIEERTSDSDSDSDS 1134
Qy      836 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 895
Db      1135 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1194
Qy      896 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 955
Db      1195 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1254
Qy      956 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1015
Db      1255 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1309
Qy      1016 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1064
Db      1310 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1359
Qy      1065 KGTLLGTLFAGIGALL--GKRKNR 1089
Db      1360 NATLFGLFALGSLTLFGRKKQNK 1385

RESULT 5
hypoetical protein sdrc [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89824
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani, U.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: C89824
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-953 <KUR>
A:Cross-references: GB:BA000018; PID:913700453; PIDN:BA041750.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: sdrc

Query Match      37.1%; Score 2093.5; DB 2; Length 953;
Best Local Similarity 43.8%; Pred. No. 5.9e-74;
Matches 488; Conservative 140; Mismatches 298; Indels 189; Gaps 28;

Qy      2  INKNNLLTKKKPIANKSKNYAIRKFTVGTASIVGATLLGIGHNEAKAEKNSYQVYKD 61
Db      1  MNKKTAATNRKGMIPNRLKFSIRKYSVGTASILVGTLLIFGLSGHEKAAE----- 52
Qy      62 SNTDELSNDOSDEKKNVYINNOSINTDNNQIIEKETNNYDIEKRSDETEST 121
Db      53 -HTNGELNOSKNETAPSEN----- 74
Qy      122 TNVDENEAFTLOKTPQDNTHLLEEVEKSSSVSSSSSIDTAQPSHTINREESVQSD 181
Db      75 EKVDSROL-----KDNQTA-----TADQPKVY----- 97
Qy      182 NVEDSHVDFANSKIKESNTESGKEENTIEOPNKKKEDSTTSQPSGYINDEK-----ISNQ 238
Db      98 -----KSDATVAKETSSNMOS-----PQNTASQSTQTSNVTNKSSTTYSENE 142
Qy      239 DELNLPINEYEN-KARPLSTSAOPSTIKRVYNOLAA-EGGSNNHLIKYTDOST----- 292
Db      143 TDKSNL--TQAKNNTSTTKTTTIKORALNRKAVNTVAAPQGTINNDKVVHTNIDIAIDK 200
Qy      293 -----TEGYDSEGVIAKHADENLIDVTEFVDOKVSGDTMVDIDKNVTPSDLT 343
Db      201 GHVKKTTGNTFEMATSSDVLK-----LKANTYIDSVKGEQTFPKYQYFRAPGSVR 252
Qy      344 DSFTIPKIKDNGELIATGYDNKNKQITTEFTDYVDKYEKNIKAKHLKLSYIDKSKVFN 403

```

```

Db 253 LPSOTNLNAGNITAKITIOSKNTTITTYTNTYDQITNVSQSFQVAFAREANATD 312
Qy 404 NTKLDEYKTAALSSVKKITTEVQRPENRNTANLQSMFNIDRKHITVEQIYIN-PLRY 462
Db 313 KRAYKMEVTLGNDYTSKDIYDY--GNQKGOQLISTSTVNNEDLSRMNVTYVNOPKRT 369
Qy 463 SAKETIV-NISGNDGSGTIIIDSTIIKYYKGDNONLSDNRITVSEVEDYND-DVA 520
Db 370 YKTEFVTMLT-----GYKFNPAKMKIYEVLDOQNFVDS-FTPTDSKIDLTGQFDTY 423
Qy 521 QGNNNDVAINGN-----IDSPILIKISKYDKNDDYTTIOOTYMQMTINYTEGPR 576
Db 424 YSNDKNTATVDLNGSSSSDKOTIIQOVAVPDSSTONGKIDYTLTQNGKSSWSN---- 479
Qy 577 ASYDNTIAFSTSGOGGLPPEKTKIGDYVEDVDKDGIONTNNENRPLSNVLTLY 636
Db 480 -SYSNNGSTANG-----QKYNLGDYVEDINKGKQDA--NEKGIRKVVYILKD 529
Qy 637 PDGTS-KSVRTDDEGKYOFDGLKNGLYKITFETPEGYPTLKHSOTNPALDSEGSVMY 695
Db 530 SNGKELDRFTDENGKYOFTGLSNG--TYSEFETPAGYPTTANAGTDADVDSGLTTTG 588
Qy 696 TINGODDMTIDSGEYOTPKYSLGNVYVDPNKGICGDDKIGSYKVTLEKDNENISTP 755
Db 589 VIKDADNMTLDGFKYTPKYSLSGDIYVYDKNKGKODSTEGIKGVKVTLOKNGEYIGT 648
Qy 756 TTTDENGKYOFTNLNSGNTVHFDPKPSGMYTQTTDSGDDDEODADEEVAVITTDHDFS 815
Db 649 TENDENGKTRFDLDSGKRVIFEKPAQLGTQNTTDD-KDADGGEVYVITTDHDF 707
Qy 816 IDNGYVDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db 708 LNGYVEEET-SPSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 766
Qy 876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db 767 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 826
Qy 936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db 827 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 886
Qy 996 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1055
Db 887 DSDSDSDAGKHPTPKPMSTV-----KDHKTAKALPE 918
Qy 1056 TGANEDYSGKTLTGLTFGLGALL-GKRRKMRK 1089
Db 919 TGSNNNSNNGTLFGSLFALGSLLEGRKKONK 953

RESULT 6
T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28679
R:Accession, E: McCrear, K.; Ni Eidhin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of St
A:Reference number: Z20510; M0ID:99098700; PMID:9884231
A:Accession: T28679
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA06651.1
C:Genetics:
A:Gene: sdid

Query Match 33.3%; Score 1880; DB 2; Length 1315;
Best Local Similarity 36.0%; Pred. No. 1 6e-65;
Matches 498; Conservative 177; Mismatches 346; Indels 362; Gaps 46;

```

```

Qy 1 MINKN-NLTKKKPIANKSNKAIKFTVGTASTIYCATLFLGLGNKAKAENSVDY 59
Db 1 MLNREKTAITRRGAWSNLNKFSIRKYVGTASILVGLTFLGLNOEKAAESTNKL 60
Qy 60 KOSNTDELDSDSDSDDEKNDVINNNOSINTDNNQIIEETNNYDIEKRSRDE 119
Db 61 NEAT-----SASNOSSD--KYDMOQNOEDNKNND--KEM-----VSSQNETTS 105
Qy 120 STTNVDENATFLQKTPODNTLJEEVEKSSSESSSSSIDTAQPSHTTINRESVOT 179
Db 106 NGKLLIEKS--VQSTGNKVEVSTAKSDQASPKSTNEDLNKQ-----TISNOALQ 157
Qy 180 SDVVEDSHVSDFAANSIKESNTESGKEENTIEOPNKYKEDSTISQPSGTYNT-DEKISQ 238
Db 158 PDLOEKRSYVN-----VQPTNEENKRYD-----AKESTT-----LNVKSDAIKSN 198
Qy 239 DELL--NLPINEYENKARELSTSAQ-----SIRKRVNOLAEOGSNVHLKVDOSI 292
Db 199 DETLVNNSNNSNNENNADITLPKSTAPKRLNTMRJAANVQPSSTEAKNVNDLITSNTLT 258
Qy 293 TEGYDSEGVAKAHDAENLIYDTFEVDKVKSGDTMYVDIDKNV-----PSDLTSP 346
Db 259 VVDADKNNKIVPAQDYLSLKSQIT--VDOKVKSGDYFTIKY-SDTYQVYGLNPEIDKN-- 313
Qy 347 TTPKID-NSGELIATGYDNKKQITTYTVDYKYEINKAKLSTYIDKSKVNNNT 405
Db 314 -IDIDDPNNGETIAAKHDTANLLITPTTVDVDRNSVQOMINSIYMDADTIP--VS 370
Qy 406 KLVYEKTAALSSVKKITTEVQRPN--ENRTANLQSMFNIDRKHITVEQIYIN-PLRY 455
Db 371 KNDYERVNTGNTTTTNTTANIQIPDYVNEKNSIGSFT--EYVSHVGNKENGKYKOTI 428
Qy 456 YINPLRSAKETVNI-----SGNGDESGTIIIDSTIIKYYKGDNONLSDNRITVSE 509
Db 429 YVNPSESLNAKLKQAVYHSSYPNNIGQINKDVTIDIKIYQVKGYL--NGGYDNTK 485
Qy 510 EYEDYTDVDAQ--LGNNDVINENIDSPITIKYISKYDKNDDYTTIOOTYMQMT 566
Db 486 ELTDVYN-QYLOKITTYGDNNSAVIDEGNADSAVVMYNTFOYTNSSPPLVGMATLST 544
Qy 567 INEYTGFTASYNMTIAFTSSGOGGGLPPEKTKIGDYVEDVDKDGIONTNNENRPL 618
Db 545 GN-----KSVSTGNALGFTNNOSGAG--QEVYKIGVYVEDTKNGVQELGKGVG 594
Qy 619 -----NTN----- 621
Db 595 NNTVTVEDNNFTNKVGEAVTKEDGSYLIPNLPGNDYVEPSNLPKGYEVRPSKQNNEL 654
Qy 622 -----DN-----EKPLSNVLTLY 636
Db 655 DSNGLSSVITVNGKDNLSADLGIYKPKYNLGDYVWEDTNKNGIODDEKIGSVTYVLKD 714
Qy 637 PDG-TSKSVRTDDEGKYOFDGLKNGLYKITFETPEGYPTLKHSOTNPALDSEGSVMY 695
Db 715 ENGLVLTVTTDADGKTKKFDLNG-NYKVEFTPEBTPTTYTSGSDIEKDSGLTTTG 773
Qy 696 TINGODDMTIDSGEYOTPKYSLGNVYVDPNKGICGDDDEODADEEVAVITTDHDFS 755
Db 774 VINGADNMTLDGFKYTPKYSLSGDIYVYDKNKGKODSTEGIKGVKVTLOKNGEYIGT 833
Qy 756 TTTDENGKYOFTNLNSGNTVHFDPKPSGMYTQTTDSGDDDEODADEEVAVITTDHDFS 815
Db 834 TKTDGKGYOFTGLNENGYVEFEETPEGYPTVQSGTDEGIDSNSTGTGVIKDKDNDT 893
Qy 816 IDNGYVDE-----SDSDSDSDSD----- 836
Db 894 IDSGFKPTYNLGDYVWEDTNKNGVQDKDEKIGSVTYVLTKDENOKYLKVTTDENGKY 933
Qy 837 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS----- 861
Db 954 FTDLNNGTYKVEFETPSGMYPTVTSNGDTEKSNGLITTGIVIKDADNMTLDGSKFYTPK 1013

```

RESULT 7  
D89852  
fibrinogen-binding protein A, clumping factor [imported] - *Staphylococcus aureus* (strain  
C:Species: *Staphylococcus aureus*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: D89852  
R:Kunoda, M.; Ohts, T.; Uchiyama, I.; Baba, T.; Yucawa, H.; Kobayashi, I.; Cui, L.; Ogawa,  
ma, A.; Mutsaers, I., Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsutsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: D89852  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1798 <KUR>  
A:Cross-references: GB:BA000018, PTD:913700678, PIDN:BA841975.1, GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
C:Gene: cflA

QY 361 TGTNRKKKQITVFTFDYVDKVENIKAKHLKLSYIDKSKVPNNNTKL-DVEKTKL--SS 417  
Db 306 NGVIDS -DGNWYITFTDIDYDKENKENTANITMPAYID--PENVTKGWVLTITGINT 360  
QY 418 VNKITVEYORPENRRTLNQSMFTNIDTKNHTVQIITVINEPLRSKAKETNV-----N 470  
Db 361 ASKTYVLIDYEKGQGHNLISIKGTIDQIDKTNNTNYQITIVNP-----SGDNVVLPALGN 415  
QY 471 ISGNDESGTIIIDSTIIKVKYKVDNQNLPSNRIRYDSEYEDVTNDYIPLQGLNNNDVNI 530  
Db 416 LIPNKRNALIDAKNTDITKYRV-DNANDLSESYVNPSEDEVTNQIRISPPANQYKV 474  
QY 531 NF----GNIDSPYIIKVIKSKYDPNKNDDYTIIQVTFMOTITNEYGEE--PRASVDNTA 584  
Db 475 EFPDDDDOITTPYIVVNGHIDPASTG-----DLALSTYIGDINSNFIWMSMDNEVA 528  
QY 585 FSTSSGQOG-DLP--PEKTYKIGDY--VWBDVDKDGIQNTNDNEKPLSHVLYLTLPYDPG 639  
Db 529 FNNGSGSDGIDKPVFPQDPDEGEIEPIPEDSD-----PGS 567  
QY 640 TSKSVRTBEDDKIQPDGLKNGLTYYITETPEGYIPTLKHSGTNPALDSKNSVWYTING 699  
Db 568 DSGSDSNSDSS-----DSGSDSTSDSGSDSASDSDSAS----- 601  
QY 700 QDDMTIDSGFYQPKYSLGNYVWYDTNKKDGIQGDDEKISGKVKTLKENGNIISTTTD 759  
Db 602 DSDKASDS-----DSASDSBASDSAS-----DSASDSDSASD 638  
QY 760 ENGRKYQFDNLNSGNYIVHFKRPSGNTQTTTDSGDDDEQADAGEEYVHTITPDHDFSIDNG 819  
Db 639 SDSASDSASDSASDSASDSASD-SDSASDSASDSASD-----SDSDSDSDSDS 689  
QY 820 YYDESDS 879  
Db 690 DS 749  
QY 880 DS 939  
Db 750 DS 809  
QY 940 DS 999  
Db 810 DS 869  
QY 1000 DS 1046  
Db 870 DS 929  
QY 1047 -----KSTKDKLPTGANEDYGSKGTLLGTLFAGLALLGKRRKNNK 1090  
Db 930 PNSPRNGTNAKNKNEAKSKPELPDTSDE--ANTSLIWGLASLGSLLFRKKENKD 987  
QY 1091 K 1091  
Db 988 K 988

RESULT 8  
S41539  
fibrinogen-binding protein - Staphylococcus aureus  
N:Alternate names: clumping factor  
C:Species: Staphylococcus aureus  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S41539; S36630  
R:McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.  
Mol. Microbiol. 11, 237-248, 1994  
A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of S.  
A:Reference number: S41539; MUID:94224142; PMID:8170386  
A:Accession: S41539  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-933 <MCD>



Db 465 RIY-----YE-----HPNVAIKFGDITKTYVVLVEGHYDNTGNLKTQVIOENY 509  
QY 562 TMOITINEXTGEFRASYDNMTAFSTSSGQGGDLPEPEKTYKIGDYWEDVDKDIQNTN 621  
Db 510 -----DPVTNRDYSI--F3W-----NN 524  
QY 622 DNEKPLSNVLTLYLPDGTSKSVRTDEDGKYQDGLKNGLYKITFETPEGYTPPLKHS 681  
Db 525 EN-----VVRYGCSA-----DSAVN-----PKDTP-----548  
QY 682 TNPALDSEGNVWVTINQODMTIDSGFYQTPKYSLGNTWIDTKDKGIQGDDEKIGISV 741  
Db 549 -GPPVDPPEPSP-----DPE-----561  
QY 742 KYTLKDEGNIIISTTTDENGKRYQFNLNSGNYIVHFKPSCGMOQTDTSDSGDDQDADG 801  
Db 562 -----PEPTPDPEPSPDPEPSPDPEP 584  
QY 802 EEVHTYTDHDFSIDNGYDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861  
585 -----DSGSDSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGSDSGS 624  
QY 862 DSD 921  
Db 625 DSD 684  
QY 922 DSD 981  
Db 685 DSD 744  
QY 982 DSD 1041  
Db 745 DSD 804  
QY 1042 G-----NSDKSTK-----DKLPDTGANEYSGKTLGLTFGLGAL 1079  
Db 805 RVTPEPNNEOKAPSNKGEVNSKNVSKOKHTDALPETG-DKSEMTNATLFGAMALLGSL 863  
QY 1080 LL-GRRKRNRKK 1091  
Db 864 LLFRKRQDKHEK 876  
RESULT 10  
AG1671  
probable peptidoglycan bound protein (LPXTG motif) lln1313 [imported] - *Listeria innocua*  
C:Species: *Listeria innocua*  
Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
Accession: AG1671  
A:Glaser, P.; Franzen, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H. Science 294, 849-852, 2001  
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1671  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-882 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC97143.1; PID:916414414; GSPDB:GN00178  
A:Experimental source: strain C1p11262  
C:Genetics:  
A:Gene: lln1313  
Query Match 13.9%; Score 785; DB 2; Length 882;  
Best Local Similarity 35.1%; Pred. No. 2,1e-23;  
Matches 151; Conservative 130; Mismatches 91; Indels 58; Gaps 9;  
QY 662 TYKITEFPEGYTPPLKHSGR-----NPAIDSEGNVWVTINQODMTI 705  
Db 95 TYTFPELPELSELTSLSPNIRKANTIDYKIAVLGIGNIGLENG-----TVNGSS-----145

QY 706 DSGFYQTPKYSLGNTVWYDNTKDGIOGDEKIGSVK-----VTLKDENGNI 752  
Db 146 NNFIDTSSNAIGARY---NHLLGV-----GVSSVITTLTIDLLALGVTLPSANDGK 196  
QY 753 ISTTTDENGKRYQFNLNSGNYIVHFKPSCGMOQTDTSDSGD-DEQDADGEVHTYTDH 811;  
Db 197 LDFAARTGGGLLDVLLNSN-----AARGFITTVGDADADADADADADADADADAD 247  
QY 812 DDFSIDNGYTDDESD 871  
Db 248 DAD 307  
QY 872 DSD 931  
Db 308 DAD 367  
QY 932 DSD 991  
Db 368 DAD 427  
QY 992 DSD 1051  
Db 428 DAD 486  
QY 1052 KLPORTGANE 1061  
Db 487 ADADADADAD 496  
RESULT 11  
AG1299  
probable peptidoglycan bound protein (LPXTG motif) lmo1799 [imported] - *Listeria mono*  
C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
Accession: AG1299  
A:Glaser, P.; Franzen, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H. Science 294, 849-852, 2001  
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1299  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-903 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC9877.1; PID:916411253; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo1799  
Query Match 13.9%; Score 782.5; DB 2; Length 903;  
Best Local Similarity 35.4%; Pred. No. 2,7e-23;  
Matches 154; Conservative 131; Mismatches 83; Indels 67; Gaps 10;  
QY 662 TYKITEFPEGYTPPLKHSNTNPALDSEGNVWVTINQODMTIDSGFYQTPKYSLGNTY 721  
Db 94 TYTFPELP-----SELSSLSNPINIRANKID---YKIAVLGIGIG 133  
QY 722 WYDTRKKDIOG-----DDEKIGSVKVT--LNDENGNIISTTTDE-----760  
Db 134 LF--NQTIVNGSSSSNFTIDSSRNAIGAKVNHLLDVGVSFTFTLTIDLLALGYTALPS 191  
QY 761 --NGRYQF-----DNLNSGNTYIVHFKPSCGMOQTDTSDSGD-DDQDADGEVHTY 806  
Db 192 ANDGKLDPFAARTGDLIDVLLNSN-----AARGFITTVGDADADADADADADADAD 242  
QY 807 TTTDHFSDIDNGYTDDESD 866  
Db 243 AD 302



[illegible]

RESULT 12

hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: F90073  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cut, L.; Ogucma, A.; Mitsuhashi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: F90073  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2271 <R>  
A:Cross-references: CB:BA000018; PTD:913702612; PIDN:BA843752.1; GSPDB:GN00149  
A:Experimental source: Strain N315  
C:Genetics:  
C:Gene: SA2447

Query Match	11.1%;	Score 624.5;	DB 2;	Length 2271;
Best Local Similarity	25.6%;	Pred. No. 1.1e-16;		
Matches 296;	Conservative 171;	Mismatches 491;	Indels 197;	Gaps 36

[illegible]

```

0Y      489 KVVYGVGNQNL-----PQSNRIYDYSEVEDYTN--D-YAQLGNNNDVNIIFNGNIDSPYII 541
Db      1632 RT-STDSQSMSLSTSTSTSMSBDSLSDSVSDSTSDSTASSTSGMSVSIISLSDSTSTS 1690
QY      542 KVISK-YDPNKDDYTTIOQTIVTMOTTINEYTGFEERTASYNTIAFSTSSGOGGODLPPEK 600
Db      1691 TSASEVMASATSDSQSMSESVDNESVESSESSEDSKMSGSTVSDDS-----GSLSYST 1745
QY      601 TYKIGDYVMEVDYDDGIQONTNDNEKPLSNLVLTITTYDDGTSGSVRT---DEDKIYOEDGL 657
Db      1746 SLRKSSESVESSSLSGSOSMSD-----SVTSDSSLSLVSJLSKSESSESDSL 1795
QY      658 KNGLTYKITFEETPEGYPTTLKHSG---TNPALD-----SEGNSYVWYINQODMTIDSGF 709
Db      1796 SDS-----KSTGSGNST-STGSLSTSTSLSGSESSESTSLSDSIISMDSSTSTSDD 1847
QY      710 YQTRKYSLGNYVWTDTKNDGICQDDEKGIGCVK-VTLKDENGNIISTTTTDENGKIYFPDN 768
Db      1848 SLSGISLISGSTSLSTS-DSSL-SDSKSLSSQSQMSGEESTSTSVSDSOSSSTSNQOFD 1904
QY      769 L-----NSGNYIVHFDKPSGMQOTTTDSGDDBDQADGBEVWYTIYDHDFG 815
Db      1905 MSIASASEDSMSTSDSSNI-----SGSNSTSTSLSTDMS-SGSYSVSTSTSLSDSIS 1956
QY      816 IDNGYVDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db      1957 GSTSVSDSSSTSTSTSLSDSMSQSQSTSTSAAGSLSTISITSTMMSMASASTSSOSTSVSTG 2016
QY      876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db      2017 LSTSDSISDSTSTISISGSGSTVEESESTSTDSISELSTSDSTSTSTSDTS6GSTS 2076
QY      936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db      2077 TSIESLSTSCSGSTSVSDSTSMGESPDSTSYMGOQDKSDSTSIDSESVSTSTSLSTS 2136
QY      996 DSDSDSDSDSDSDSDSVSDSDSDSDSDSDSDSD----- 1032
Db      2137 DSTSTSESLSTMGSQSIISDSTSTSMGGSTSTESNSMPDSDMSMHHTSTSTSLSS 2196
QY      1033 ---SDSNDSDLG-----NSSDKSTKDLRPDTGANEYGSKGTLLG---TLFAGLG 1077
Db      2197 EATSTSESQSTLATSEVTAKHNGTPAQSEKRRLPDTG---DSIKONGILGGLVMTLLVLGLG 2253
QY      1078 ALLGKKRRKNKKN 1092
Db      2254 LM---:::~::~: 2264

```

## RESULT 13

SRP40 protein from yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein YKR092c; protein YKR12a  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 03-May-1994 #sequence, revision 03-May-1994 #text, change 21-Jul-2000  
C:Accession: S38170; S40645; S37702  
R:Balladron, V.; Ballista, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantalejo, J.; Remacha, M.; submitted to the Protein Sequence Database, March 1994  
A:Reference number: S38158  
A:Accession: S38170  
A:Molecule type: DNA  
A:Residues: 1-406 <BAL>  
A:Cross-references: EMBL:Z28317; NID:G486580; PID:G486581; MIPS:YKR092c  
A:Experimental source: strain S288C  
R:Bou, G.; Esteban, P.F.; Balladron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; yeast 9, 1349-1354, 1993  
A:Title: The complete sequence of a 15 820 bp segment of *Saccharomyces cerevisiae* chr.  
A:Reference number: S40644; MUID:94205265; PMID:8154186  
A:Accession: S40645  
A:Molecule type: DNA  
A:Residues: 1-406 <BOU>  
A:Cross-references: EMBL:X73541; NID:G450550; PIDN:CA51946.1; PID:G450552





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 14.3513 Seconds

(without alignments)  
3155.972 Million cell updates/sec

Title: US-09-147-405B-15

Perfect score: 5646  
Sequence: 1 MINKNNLTKRKPIANKSN.....FAGLGALLGKRRKRNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1319.5	23.4	1637	1	MRSP_STAU P80544 staphylococ
2	708.5	12.5	1253	1	DSP_HUMAN Q9N2W4 homo sapien
3	675.5	12.0	934	1	DSP_MOUSE P97399 mus musculi
4	606.5	10.7	687	1	DSP_RAT O62598 rattus norv
5	554	9.8	406	1	SR40_YEAST P32583 saccharomyc
6	487.5	8.6	1018	1	FNBA_STAU P14738 staphylococ
7	471	8.3	640	1	SANT_PLAFV P13621 plasmodium
8	388.5	6.9	695	1	ARP_EUGR 004732 euglena gra
9	368	6.5	1664	1	SLPL_CLOTM 006852 clostridium
10	361.5	6.4	400	1	RTOA_DICDI 003681 dictyostell
11	343.5	6.1	1658	1	YMG7_YEAST 001033 herpesvirus
12	333	5.9	797	1	VG48_HSVSA 013368 candida alb
13	331.5	5.9	1419	1	ALAI_CANAL P46591 candida alb
14	315	5.6	937	1	HYRI_CANAL P14248 plasmodium
15	311	5.5	2452	1	RPB1_PLAFD P47035 saccharomyc
16	310	5.5	1189	1	YH6_YEAST P46100 homo sapien
17	303	5.4	2492	1	ATRX_HUMAN O55188 mus musculi
18	301.5	5.3	503	1	DMP1_MOUSE P98193 rattus norv
19	295.5	5.2	489	1	DMP1_RAT P18899 saccharomyc
20	294.5	5.2	429	1	DR48_YEAST 001550 xenopus lae
21	291	5.2	1744	1	TANA_YENLA P07866 bombyx mori
22	288	5.1	389	1	SER1_BOMMO 013316 homo sapien
23	285.5	5.1	513	1	DMP1_HUMAN P39653 streptococ
24	283.5	5.0	1337	1	DEXT_STRDO P39793 bacillus su
25	283	5.0	914	1	PBPA_BACSU P39520 saccharomyc
26	276.5	4.9	1085	1	IFH1_YEAST P41891 schizosacch
27	276	4.9	500	1	GAR2_SCHPO 012176 saccharomyc
28	271.5	4.8	1025	1	MK21_YEAST O61667 mus musculi
29	271	4.8	2476	1	ATRX_MOUSE 009564 caenorhabdi
30	270	4.8	688	1	LIP_STAEP 002510 staphylococ
31	269.5	4.8	3178	1	YS89_CAEEL P22290 plasmodium
32	267.5	4.7	1070	1	PVDR_PLAVS P87498 gallus gall
33	266.5	4.7	1912	1	VIT1_CHICK

34	265	4.7	593	1	SANT_PLAF7	Q03400 plasmodium
35	264.5	4.7	667	1	CYL1_BOVIN	P35662 bos taurus
36	264.5	4.7	936	1	FH11_YEAST	P39521 saccharomyc
37	264	4.7	918	1	YMBJ_CAEEL	P34487 caenorhabdi
38	263.5	4.7	1770	1	PMCR_CHLTR	O84419 chlamydia t
39	262.5	4.6	279	1	SPBP_RAT	P08723 rattus norv
40	262.5	4.6	817	1	YGA4_YEAST	P46949 saccharomyc
41	262	4.6	510	1	DMP1_BOVIN	Q95120 bos taurus
42	260.5	4.6	852	1	SRCH_RABIT	P16230 oryctolagus
43	260	4.6	1153	1	PVDR_PLAKN	P50493 plasmodium
44	260	4.6	1183	1	CNA_STAU	O53654 staphylococ
45	259.5	4.6	1435	1	EBAL_PLAF6	P19214 plasmodium

## ALIGNMENTS

RESULT 1  
MRSP\_STAU STANDARD; PRT; 1637 AA.  
ID MRSP\_STAU  
AC P80544; O92F62;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Methicillin-resistant surface protein precursor.  
GN PLS.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=Isolate 1061;  
RA Savolainen K., Kuusela P., Paulin L., Korhonen T.K.;  
RT "Pls. a large repeat-rich surface protein of methicillin resistant  
RT Staphylococcus aureus.";  
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
RN (2)  
RP SEQUENCE OF 624-628; 676-682; 938-948; 1156-1168; 1176-1185;  
RP 1199-1205 AND 1217-1224.  
RC STRAIN=Isolate 1061;  
RX MEDLINE=96270743; PubMed=8665912;  
RA Hilden P., Savolainen K., Tynneiae J., Vuento M., Kuusela P.;  
RT "Purification and characterisation of a plasmid-sensitive surface  
RT protein of Staphylococcus aureus.";  
RL Eur. J. Biochem. 236:904-910(1996).  
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
CC an amide bond (Potential).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF115379; AAD09131.1;  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR TIGRFAMs: TIGR01167; LPTG\_anchor; 1.  
DR TIGRFAMs: TIGR01168; YSTR\_signal; 1.  
DR PROSITE: PS50847; GRAM\_POS\_ANCHORING; 1.  
KW Antibiotic resistance; Glycoprotein; Cell wall; Peptidoglycan-anchor;  
KW Repeat; Signal.  
FT SIGNAL 48  
FT CHAIN 1 1601 METHICILLIN-RESISTANT SURFACE PROTEIN.  
FT PROPEP 1302 1637 REMOVED BY SORTASE (POTENTIAL);  
FT DOMAIN 1601 1582 141 X 2 AA TANDEM REPEATS OF D-[SAG].  
FT SITE 1598 1602 LPTG SORTING SIGNAL (POTENTIAL).  
FT MOD\_RES 1601 1601 AMIDE-LINKED TO CELL WALL (POTENTIAL).  
SQ SEQUENCE 1637 AA; 174573 MW; 75BE9ADB469B309 CRC64;  
Query Match 23.4%; Score 1319.5; DB 1; Length 1637;

Best Local Similarity 27.5%; Pred. No. 5.9e-43;  
Matches 461; Conservative 180; Mismatches 401; Indels 635; Gaps 60;

```

QY 7 NLLTKKP--IANKSKAIAKRFVGTASIVIGATLLFGLGHNKAKEENSVDQYKDSNT 64
D 2 NNSKKILDFLPNKIKYKISIRFVTGASIIIVGATLIFGVANDQAEAMENNTTQKODSS 61
QY 65 D-----DELSDSDSSDEKKNV INNOSINTDDNQ 97
D 62 DASKVGVNQTIEGSSANSNESDIPQVDVTKDTTEQASTEEKATTEQASTEEKADTTE 121
QY 98 IIRKEETNNVDIEK-----RSEDRTESTTNNVDENATEFLQ 133
D 122 QATTEAKAEKEDTKVEFEAKAEETDKATEEAPKAEETDKATEEAPKAEETDKATEE 181
QY 134 KTP---QDNHLLTEEVSESSSVSSSIDTAQOPSHTTINREESVOTSDNV----- 183
D 182 EAPAEETSKATEEAPAEETSKATEEAPAEETSKATEEAPAEETSKATEEAPAEETSKATEE 241
D 184 --EDSHVDFANSKIKESNTESGKEENTIEQPNKYKEDST-----TSQPS--G 227
D 242 AETSKATEEAPKAEETNKVETEAPAEETNKATEETPAVEDTNAKSNSNAOPSETE 301
QY 228 YTNIDEKISNDELINLPINEYE---NKARPLSTTS--AQPISIKRVYVQLAEOGSNVN 282
D 302 RTQVVDVTA-KDLYKKSEVTEAKAEIEKVLPKDISNLSNEEIKKIALSEVLKFAKKEN 360
QY 283 HLKATYDOSTIEGYDSEGVKAKHAENLIYDT---FEVDKVKSGD---TMTVDIDK 335
D 361 AOPRAETFSVSNNARTT-----NVNYSATLRAAODTVTKKGNGFTAGDIIH 410
QY 336 NTVPSDLDSFTPIKINSGEIATGT-----YDNK---NKQITYT----- 374
D 411 KTYKEEFNEGLTLPNTNPN--NTGTGALEYNDKIDFNKDFITIVPVANNOCNTTG 468
QY 375 -----FT-----DYVDKENIKAKLTSYIDSKVNNNTKLOVEYKFALESVNTI 422
D 469 ADGWFMFTQGGQDFLNOGILR-----DKGMANASFKIDTAYNNVNGVKDKL- 518
QY 423 TVEYORPNEKNTANLQSM-----FTN-----IDTKNHTVEQTIYIN- 458
D 519 -----DADKTNLSQIGAAGVGYGTGVKNCAGVTNOVGNALNTKDKPKNKIIYADN 571
QY 459 -----PLRYSAKETNV----- 469
D 572 TTNHLDGQFHGORLNDVYLNDAATSTITATYAGTKWAKATDGLGIDKSOKYFELITSH 631
QY 470 --NISGND-----EGSTI-----IDSTLIK-----YK 492
D 632 MONKRSNKGIMRTNLEGVITTTPOADLIDDEVTKQPIFKHKTITREDDPLLEPGSPVIVOK 691
QY 493 VGDNONL-----PDSNRIVDYSE-YEDVYNDQYAL-----GNNNDVYNIN- 532
D 692 GEDGKTTTTPKVDPRGDVVERGEPTTEVTKNVDIVHTPREVPQGHKDEDFPMLP 751
QY 533 -----GNISPIYIKYISKDPKKDYT----- 555
D 752 IDGTEEVGKPGIKNPETGEVVT--PRVDVTKHGPKAGEBEVTKKEIPEKKKEFNPD 808
QY 556 -----TFOQVTMOTIN----- 568
D 809 LKPGGEKATQEGOGKEKTTTPTTINPLTGKVGSGEPTTEVTKAPVDELITQFGGEEVPO 868
QY 569 ---EY-----TGERTASVYNTIAFSTSSGOG----- 593
D 869 GHKDFDNLPIDTEEVGKPGIKNPETGEVVPVPRVDVTKHGKKAEBEPTKKEIPE 928
QY 594 ---GLPP-----EXT-----YKIG-----DIWEDVD----- 613
D 929 KKREPNPLKGEKEKVTQEGOTGKTTTPTTINPLTGKVGSGEPTTEVTKKEPVDITQ 988
QY 614 -----KDGIONTNDN-----KPLSNVLYLTITPDGSGK 642

```

```

D 989 FCGEEVPOGHKDEFPDNLPIDTEEVGKPGIKNPETGEVVPVPRDD--VTYHGKAGEP 1046
QY 643 SVRTDE---DGKQPD-GIKNGLYTKITFEPEGYTPTLKHSCTPALD---SEGN-SVM 694
D 1047 EYTKKEIPEKKEKREPNPLKPG-EKVYQEGOTGKTTTPTTINPLTGKVGSGEPTTE 1105
QY 695 VTINQODMTIDSGFYQTPKYSLGNYVTDN-----KQGIQ----- 732
D 1106 VKPEPVDIT-QFGGEEVPO---GHKDEFPDNLPIDTEEVGKPGIKNPETGEVVPV 1161
QY 733 DD-----EKGISGVKYVLKDNNGNIISTTT----- 758
D 1162 DVTYHGKAGEBEVTKKEIPEYTRVLDPTIMEPG-SPKVAQKGBNGEKTTTPTTINP 1220
QY 759 -----DENGKY-----QPD-NLNSG----- 772
D 1221 LTGEKVGGEPTTEVTKPEIDIVNYAPRIPIHGTRIEDPMLPGETGVYIGKDGKLP 1280
QY 773 -----NYIYHFKPSGMTOTTTDSCGDDQDADAGEVHYTITDHDFSIDNGY 821
D 1281 ETGEIIEEPQDEVIITHGAKDSDADSDADSDADSDADSDADSDADSDADSDADSDADSD 1340
QY 822 DDESDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 881
D 1341 DSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1400
QY 882 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 941
D 1401 DSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1460
QY 942 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1001
D 1461 DSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1520
QY 1002 DSDSDS-----DSDSVSPSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1035
D 1521 DSDSDADSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1580
QY 1036 DNDSDLGNSDSKSTKDKLPDTGANEDYSGKTLGLTFAGLAL-LLCKRRKRRNK 1091
D 1581 DADRDHNDKTKDPKNKELPDTC--NDAQNNGTLFQSLFALGLGLVYGRRRKKNNE 1635

RESULT 2
DSDP_HUMAN
ID DSDP_HUMAN STANDARD; PRT; 1253 AA.
AC O9NZM4; O95815;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Dentin stialosphoprotein precursor [Contains: Dentin phosphoprotein
DE (Dentin phosphoprotein) (DPP); Dentin stialoprotein (DSP)].
GN DSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20168992; PubMed-10706475;
RA Gu K., Chang S.R., Ritchie H.H., Clarkson B.H., Rutherford R.B.;
RT "Molecular cloning of a human dentin stialosphoprotein gene.";
RL Eur. J. Oral Sci. 108:35-42(2000).
RN [2]
RP SEQUENCE OF 463-1253 FROM N.A.
RX MEDLINE-99094526; PubMed-9879917;
RA Gu K., Chang S.R., Slaven M.S., Clarkson B.H., Rutherford R.B.,
RA Ritchie H.H.;
RT "Human dentin phosphoprotein nucleotide and amino acid sequence.";
RL Eur. J. Oral Sci. 106:1043-1047(1998).
RN [3]
RP DISEASE.

```





Db 214 GISENGTEFTPSIGDADGDDTDGSDSGNGVEDEDTGSGD----- 256  
 QY 257 STTSAPSIKRV---TVNOLAEQGSNVHLIVYTDOSTEGYDSDS---EGVIKADAE 309  
 Db 257 -GEGAEAGDGRSHDGTGGCGGSHGNTDHR---GOSVSTEDDSEKQEGFPNGHND 312  
 QY 310 NLIDVYFEVDKRYSGDTMVDIDKNV--PSDLT-----SFTPKIKDNGELIA 360  
 Db 313 N-----SSENGVEEDSGTQATQDKELSPKDRDAEGGIISGEACPSGSDQGIET 366  
 QY 361 TGTVDNK-NKOITFTDYDYKYENIKAKHLTSTYIDKSKYPNNNTKLDVEKTLASSVN 419  
 Db 367 EG--PKNKSKIT-----TKESGLSSKXSGNGHQVELDKRN----- 403  
 QY 420 KTIVYQRPENRNTANLQSMFTNIDTKNHTVEQTIYINPLRYEAKETNVNISNGDEGS 479  
 Db 404 -----PKQGES-----DKPQGTAEK-----SAAHSNLGHSRIGSSSN 435  
 QY 480 TIIDSTIKYKYKGDQNLDPDSNRIYSEYEDVTNDYAQLGNNNVINFGNIDSPY 539  
 Db 436 SDGSDS---YEPDDESMDGDDPKSDSDESGSDSDT-----NSGANESGSGDASY 484  
 QY 540 IIKYISKYDPKDDYTTIQGTVMQTIYNEYGEFRTASYDNTIAFSTSGGCGDLPPE 599  
 Db 485 TSDSSDDDDNDSGH-----AGEDD--SSDSSGGGSD----- 516  
 QY 600 KTYIGYDVEDVDKGIQNT-NDNEKPLSNVLTLYTPDGTSKSVRID--EDGKYQPDG 656  
 Db 517 ---SNGDDSDSEDEKDESDSDHNS-----DSEKSDSDSDSDSDSDSDSDSDS 561  
 QY 657 LKNGITYITETPEGYPTLKHSGTNPALDSEGSVWVTINGGDMDTDSGFYQTPKYS 716  
 Db 562 -----SDSSSDS 604  
 QY 717 LGNYVYDT-----NKDGIGQDEKIGSVKVTLKDEMN--ISTTTDNGKYYQFDN 768  
 Db 605 -----DTCCSDS 657  
 QY 769 LNSGNYIHFDPKSGMTQTTTDSGDDDEQADAGEVHYATTIDHDFSIDNGYIDDESDD 828  
 Db 658 SDSDSD-----SSD 709  
 QY 829 SD 888  
 Db 710 SSDSGSSSDS 768  
 QY 889 SD 948  
 Db 769 DS 828  
 QY 949 SD 1008  
 Db 829 DSD 888  
 QY 1009 SD 1051  
 Db 889 SKDSSSDSDSDGSKSGNGNSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 933

## RESULT 4

ID DSPP\_RAT STANDARD: PRT: 687 AA.  
 AC Q62598; Q9R057; P70578;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Dentin staphosphoprotein precursor [contains: Dentin phosphoprotein  
 (Dentin staphosphoprotein) (DPP); Dentin staphosphoprotein (DSP)].  
 GN DSP OR R0SP2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM DPP-2).  
 RC STRAIN-Sprague-Dawley;  
 RA MEDLINE=20435277; PubMed=10978503;  
 RX Ritchie H.H., Mang L.-H.;  
 RT "The presence of multiple rat DSP-PP transcripts.";  
 RL Biochim. Biophys. Acta 1493:27-32(2000).  
 RN [2]  
 RP SEQUENCE OF 1-387 FROM N.A., AND SEQUENCE OF 18-26.  
 RC STRAIN-Sprague-Dawley; TISSUE=Odontoblast;  
 RX MEDLINE=94148875; PubMed=8106414;  
 RA Ritchie H.H., Hou H., Veis A., Butler W.T.;  
 RT "Cloning and sequence determination of rat dentin staphosphoprotein, a  
 novel dentin protein.";  
 RL J. Biol. Chem. 269:3698-3702(1994).  
 RN [3]  
 RP SEQUENCE OF 421-687 FROM N.A. (ISOFORM DPP-1).  
 RC STRAIN-Sprague-Dawley;  
 RX MEDLINE=96355551; PubMed=8702961;  
 RA Ritchie H.H., Mang L.-H.;  
 RT "Sequence determination of an extremely acidic rat dentin  
 phosphoprotein.";  
 RL J. Biol. Chem. 271:21695-21698(1996).  
 RN [4]  
 RP SEQUENCE OF 1-19 FROM N.A.  
 RC STRAIN-Sprague-Dawley;  
 RX MEDLINE=99333695; PubMed=10403786;  
 RA Yamazaki H., Kunisada T., Miyamoto A., Tagaya H., Hayashi S.-I.;  
 RT "Tooth-specific expression conferred by the regulatory sequences of  
 rat dentin staphosphoprotein gene in transgenic mice.";  
 RL Biochem. Biophys. Res. Commun. 260:433-440(1999).  
 RN [5]  
 RP SEQUENCE OF 29-33; 70-79; 93-109; 136-148; 162-188; 266-308; 398-423  
 AND 426-438, AND PHOSPHORYLATION OF SER-292 AND SER-298.  
 RX MEDLINE=21125612; PubMed=11042175;  
 RA Qin C., Cook R.G., Orkiszewski R.S., Butler W.T.;  
 RT "Identification and characterization of the carboxyl-terminal region  
 of rat dentin staphosphoprotein.";  
 RL J. Biol. Chem. 276:904-909(2001).  
 RN [6]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=98055479; PubMed=9395101;  
 RA Ritchie H.H., Berry J.E., Somerman M.J., Hanks C.T., Bronckers A.L.,  
 Hottel D., Papagerakis P., Berdal A., Butler W.T.;  
 RT "Dentin staphosphoprotein (DSP) transcripts: developmentally-sustained  
 expression in odontoblasts and transient expression in  
 pre-ameloblasts.";  
 RL Eur. J. Oral Sci. 105:405-413(1997).  
 CC - FUNCTION: DSP may be an important factor in dentinogenesis. DPP  
 may bind high amount of calcium and facilitate initial  
 mineralization of dentin matrix collagen as well as regulate the  
 size and shape of the crystals.  
 CC - SUBCELLULAR LOCATION: Secreted.  
 CC - ALTERNATIVE PRODUCTS: 2 isoforms; DPP-1/PP240 (shown here), and  
 DPP-2/PP171; may be produced by alternative splicing.  
 CC - TISSUE SPECIFICITY: Specifically expressed in teeth, mainly in  
 odontoblasts and transiently in pre-ameloblasts.  
 CC - PTM: DSP is glycosylated.  
 CC - CAUTION: Ref. 2 sequence differs from that shown due to a  
 frameshift in position 380.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: AF247187; AAK96895.1;  
 DR EMBL: 002074; AAA18932.1; ALT\_FRAME.  
 DR EMBL: U63111; AAC52774.1;



DR EMBL: AF114987; AA048588.1; ALT-TERM.  
 KW Signal: Extracellular matrix; Glycoprotein; Stallo acid;  
 KM Phosphorylation; Alternative splicing.

FT SIGNAL 1 17  
 FT CHAIN 18 687  
 FT CHAIN 18 447  
 FT CHAIN 448 687  
 FT MOD\_RES 57 57  
 FT MOD\_RES 226 226  
 FT MOD\_RES 253 253  
 FT MOD\_RES 278 278  
 FT MOD\_RES 292 292  
 FT MOD\_RES 298 298  
 FT MOD\_RES 315 315  
 FT MOD\_RES 319 319  
 FT MOD\_RES 329 329  
 FT MOD\_RES 337 337  
 FT MOD\_RES 345 345  
 FT MOD\_RES 366 366  
 FT CARBOHYD 55 55  
 FT CARBOHYD 82 82  
 FT CARBOHYD 128 128  
 FT CARBOHYD 189 189  
 FT CARBOHYD 312 312  
 FT CARBOHYD 369 369  
 FT VARSPPLIC 567 635  
 FT CONFLICT 74 74  
 FT CONFLICT 564 564  
 FT SEQUENCE 687 AA: 70179 MW: 9A845ED6AA31B63 CRC64;

Query Match 10.7%; Score 606.5; DB 1; Length 687;  
 Best Local Similarity 32.08; Pred. No. 1.6e-16;  
 Matches 211; Conservative 66; Mismatches 250; Indels 131; Gaps 30;

QY 453 QTYINPLRSYAKET-----NYNISNGDEGT-IIDSTIKYKV-GDN 496  
 DB 97 QPLIANGSNAKREVEYTHGDIHAGENSTANGIRGVGLAEMAEKEKVGOP 156  
 QY 497 QNLPDSRIIDYSEYEV---NDVYALQNNNDVNFNGIDSPYITIKYISKYPRKDD 553  
 DB 157 HODTKGLASDTSONGATLVQENEPQVAGSKNSTNEVTHGSGV---AAQETTPQREG 213  
 QY 554 YTTIQVTQNTINEYTGERTASVDNTIAFTSSGOGGDDLPKTYTKGVDVWEDV- 612  
 DB 214 EGSENGAEVTPSIGEGAG-----LDNT-----EGSPSGNGIEDEDTGS-GDVGADAG 262  
 QY 613 ----DKDGIQ-----NNDNEKPLSNVLTYPDGTSKSVTDEDEKYGQDGLKNG 660  
 DB 263 DGRSHDGTGHEGSGSSGNNDR-----GGGVSFTDDDSKDEGSPNG 307  
 QY 661 LTKITFEPEGYTPILKHSCTNPALD-----SEGSVWVTINGQDDMTIDSGFYQTPKY 715  
 DB 308 -----RGCDNTSSSEETGIEBGDGTQTDQDNQNLSPREGGIIISQAEA 349  
 QY 716 SLGNYVYDTNKDIOGGDDKGIS-GVKYTLKNDNGNIITTTDDNGKY--QDNLNSG 772  
 DB 350 C-----PGQSONOGLIEGSTGKSSITYESGKL--SGSND3GHHGEMELDKRNSP 400  
 QY 773 NYIVHFKPESGNTQ---TTTDSGDD---DEQADGEEVHYITFDHFSIDNGYVQDESD 826  
 DB 401 KQ-GESEKPGGAKEKSTHNNMGHSRTGSSSSNSGHSY---DFDESMQ---DDPNS 452  
 QY 827 SDDSDSDSDSDSDSDS-----DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 874  
 DB 453 SDESNCGSDGSDSDSEBAIENGNGHGDASYTSDESSDNGSDSHHA3BDDSDPTD-D 511  
 QY 875 SDDSDSDSDSDS-DS 933  
 DB 512 SDEKGGDDSDSKDDESDNNDNDSDSEKSDS-SDDSDSDSDSDSDSDSDSDSDSDS 570  
 QY 934 SDDSD 992  
 DB 571 SDESDTSDS 629

QY 993 SDDSD 1051  
 DB 630 SDDSD 686

RESULT 5  
 SR40\_YEAST STANDARD; PRT; 406 AA.

ID SR40\_YEAST  
 AC P32583;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Suppressor protein SRP40.  
 GN SRP40 OR YKR092C OR YKR412A.  
 OS *Saccharomyces cerevisiae* (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID:4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN:ATCC 28563 / FL100;  
 RA Lalo D., Charles C., Sentenac A., Thurlaux P.,  
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.  
 RX MEDLINE:94205265; Pubmed-8154186;  
 RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,  
 RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.,  
 RT "The complete sequence of a 15,820 bp segment of *Saccharomyces*  
 RT *cerevisiae* chromosome XI contains the UBI2 and MFI1 genes and three  
 RT new open reading frames".  
 RL Yeast 9:1349-1354 (1993).  
 CC -!- FUNCTION: NOT KNOWN; WEAK SUPPRESSOR OF A MUTANT OF THE  
 CC SUBUNIT AC40 OF DNA DEPENDANT RNA POLYMERASE I AND III.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L11275; AAA35091.1;  
 DR EMBL: X73541; CAA51946.1;  
 DR EMBL: Z28317; CAA82171.1;  
 DR PIR: S38170; S38170.  
 DR SGD: S0001800; SRP40.  
 FT DOMAIN 25 314  
 FT CONFLICT 400 400 G->N (IN REF. 1).  
 FT SEQUENCE 406 AA: 41015 MW: 8EA007695AF4BAID CRC64;

Query Match 9.8%; Score 554; DB 1; Length 406;  
 Best Local Similarity 46.8%; Pred. No. 8.3e-15;  
 Matches 137; Conservative 25; Mismatches 93; Indels 38; Gaps 6;

QY 806 VTITDHDGFDSDNGYDS 865  
 DB 6 IKDVEVRLKLYKEKEIEKSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 65  
 QY 866 DSD 924  
 DB 66 DSDSDSDSES-SSSESD 124  
 QY 925 -----SD 952  
 DB 125 ETKKRARESNEDAKETKAKTEPSSSSSESSSSSSSSSSSSSSSSSSSSSSSSSSSS 184  
 QY 953 SDDS 1009  
 DB 185 SDESEDESDSSSDS 244









DB 303 SSDGKTTCTISFHDTLSINTVDDDEICTGKE 334

RESULT 11  
Y67 YEAST  
ID Y67 YEAST STANDARD: PRT: 1658 AA.  
AC 003661; 004988;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 Intergenic region.  
GN YMR219W OR YMR261.13 OR YMR959.01.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
ON NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE OF 1-711 FROM N.A.  
RC STRAIN=S288c / AB972;  
RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,  
RA Walsh S.V.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 608-1648 FROM N.A.  
RC STRAIN=S288c / AB972;  
RA Skellon J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: 249809; CA89934.1; -  
DR EMBL: 249933; CA90190.1; -  
DR SGD: S0004832; YMR219W.  
KM Hypothetical protein.  
SQ SEQUENCE 1658 AA; 187137 MW; 3893f968305A757D CRC64;

Query Match 6.18; Score 343.5; DB 1; Length 1658;  
Best Local Similarity 19.2%; Pred. No. 3.3e-06;  
Matches 241; Conservative 215; Mismatches 446; Indels 355; Gaps 57;

QY 47 NEAKAEENSVDYKSDTDELSDNDSSDE-----KKNVINNNNSINDDNNQIHK 101  
DB 232 DEEYAEAGLQDY-----SNDYAEEGOVERKN--IQEQA-NVENATQISSS 277  
QY 102 E--ETNNY-DGIEKRSERDRTSTTNVDENATFLQKTPQ-----DNT----- 140  
DB 278 DSSGQVYSGEVELEDDIDVESDAKDESGAEGTEHSD?SKYQPRDNKIPVIE 337  
QY 141 -----HLTEEVKESSVE--SSNSIDTAQOPSHTTI-----NREESYQTS 180  
DB 338 KYSEDEKRVHQRYSEDGAFGFSVNIISVDDSEDEESQAESYSAEAENVYHNEHELDK 397  
QY 181 DNVDSHVSDFPANKIKESNTESKENTIEQPKVKEDSTTSQPSGT--NIDEKISNQ 238  
DB 398 ELLEDISSSESOSAES--EAGSE--DFEYKMKNEKSTSEETNTESRQGF-K 451  
QY 239 DELLNLPINEYENKARPLSTSAQPSI-KRYTVNQLAAEQSNVNHILKIVDOSITE-- 294  
DB 452 DAYTKNVQEOENDEPEKDOIIRSSLDKFNHNNKSEYENV--LENEIDPAIVEREN 509  
QY 295 GYDSEGVYAHDAENLIYDVFEEVDKVKSGTMTYDIDKNTVPSDLTDSFTTPKIDN 354  
DB 510 QINDVEG-----YDVT-----GKSVESDLHHS--PDNLXDAAARALQFQ 548  
QY 355 SGEIIATGYDNKNKOITVPTDYVD-----KYENIKAKLKLSYDKSKVNNNTKL 407  
DB 549 QSR---NSNCPQKEEQVSESTLGHNSGNSLGRSLDSEEDQIPLKDTFGEN---NNNLKT 602

QY 408 DVEKTAALSSVKNITTYEYQRPENRPAANQSMFTNIDTNHNYEOTIYINPLRYAKET 467  
DB 603 D-----RGDLSSSVLEIEVKVSEKKLDG-----STERELY-----PL---STDY 638  
QY 468 NVNISGNGDEGRTI--RDDSTIIKVVYVGDNONLPD-----SRIIDYS 509  
DB 639 TINNSLGNEDSIYSLDDADAI-----SENLTDPPLMEIKTTPRYEVYISGSYST 691  
QY 510 EYEDVT-----NDYVLAQGNNDVYNIN-----FGNIDSPYI 540  
DB 692 SYEDNTYAMPQVEYTPSPFNNDPNSLNDDEYK---KHDLKSTLALAPAFKKDAFEV 748  
QY 541 IKYISK-----YPPNKD---DYTTIQGVMTQITNEYTGEFRASYNNTI 583  
DB 749 EAGVTAKCLTSTSGHTNIFHTSKETKQVSDLESTENVTFE---NENYGDENKQSKNFP 805  
QY 584 AFTSSGQGGDLPREKTYIKGVYVMDYDKDQGNQNDNEKPLSNVLYVLTYYD---GT 640  
DB 806 GVANSTDKSTEDNTDEYFSAINY---TNTVGSCEDEIETASVNEENLRCEKDMNE 861  
QY 641 SKSVRTDEGKYQFGDKNGLTLYKITPEPEGGYPTLKHSGTNPALDSEGSVYVITNGQ 700  
DB 862 AEMSGDECYKQNDGSKTQISF--STDSPDNF-----QESNDN---TEFSSTYKVRNS 911  
QY 701 DDMTIDSGFYQTPKYSLGNYVYDTNKGIDQ--GDEKIGSVKTYLKDE--NGNIIST 755  
DB 912 DLEDDELKKEITLKAELVADKLEDEESSEYEQVADPEPG-----NDEGSNENIVKG 963  
QY 756 TTIDENKRYQPDN-----LNSGNYIYHPKPGSMQTTIDSDDD 795  
DB 964 TKKDTLIVERENKRVKHEETLFEANVSSVYONKMDHTOVINDQANVAGEERK 1023  
QY 796 E--ODADGEEVHYVITDH--DPFSIDN-----818  
DB 1024 YIIONTPTEPHIISIERIDENALGNMEIPRSCVETKTHNEVLFERRATIENTKALEN 1083  
QY 819 --GYDDESDS 873  
DB 1084 NTNHMDQVQACSDSDRQDQSTAEKNEGSAKHLMDIRVSSSEIESVEPLPESDRSNIF 1143  
QY 874 DS-----SDS 903  
DB 1144 SSPIRIGAVVYGVKYVDAEVSFKRIDVNDSDSDNVDIGDQNOGIFNKSNS--TDASV 1202  
QY 904 DS 949  
DB 1203 NMAVSSKERSDDEDAVILGVTAAEAHNDGNSRVINIDPTNGAYEEDSEVFRQOVK 1262  
QY 950 DS 1009  
DB 1263 DKENLHKSEEPVLEGLOEQHFKEKKHSENEEEDTIVGDTISANISHNAPDDIKROQL 1322  
QY 1010 DSVSDSDSDS-----DSDSGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1061  
DB 1323 KNLSDLENYSQLRLEDSRGRKNOESEDENVTSRERDL--TFEKSVMKEY--AGAIEE 1375

RESULT 12  
VG48\_HSVSA  
ID VG48\_HSVSA STANDARD: PRT: 797 AA.  
AC 001033;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last annotation update)  
DE Hypothetical gene 48 protein.  
GN 48 OR EDLF5.  
OS Herpesvirus salm1r1 (strain 11).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Rhadinovirus.  
ON NCBI\_TaxID=10383;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE-92333688; PubMed-1321287;  
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,  
 RA Newman C., Wilmann S., Craxton M.A., Coleman H., Fleckenstein B.,  
 RA Honess R.W.;  
 RT "Primary structure of the herpesvirus saimiri genome.";  
 RT J. Virol. 66:5047-5058(1992).  
 CC -1- SIMILARITY: TO BBV BRRE2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X64346; CAA45671.1;  
 CC PIR: A36811; A36811.  
 DR Hypothetical protein  
 V SEQUENCE 797 AA: 88867 MW: 9294234AD850E23 CRC64;

Query Match 5.9%; Score 333; DB 1; Length 797;  
 Best Local Similarity 18.8%; Pred. No. 3.7e-06;  
 Matches 116; Conservative 104; Mismatches 275; Indels 122; Gaps 22;

QY 506 YQYSEEDVTDNDYQALGNNDVNFNGINDSPYIKYISKDPMKDYTTTQQTVMOT 565  
 DB 194 YQFNEK-----NEISLVATCINCWCWLFMLQOY--MSSDLALALN-RT 237.  
 QY 566 TNEYGEFRTASDNTAFSTSSGQO-----GDLPEKTKIGDYVMDVKDGI 617  
 DB 238 YLALHND--KASYSILKFLTNSHREHVTKVANKAMOSLKI-----I 283  
 QY 618 QNTNDNEKPLSNVY-----TLTP--DGTSKSVRTDEDEKQYQDGL 657  
 DB 284 KQTEKPSKTKMLMISILSGRIGMDLFCQSVLKAPLIDHKLSPVSEYED--FD-- 337  
 QY 658 KNGLYKTIFFPEEGTPTLKHSGTNPALDSGNSY-----WTINQD----- 701  
 DB 338 DEVELCISDDEVD-----SEDGMLCVLDDSESVNVALROVLVVDKQANEKYEKII 391  
 QY 702 DMTIDSGFYQTKYSIGNVYVDTNKGIGQDEKGISGVKTYLNDENNIITTDEN 761  
 DB 392 DKSDDRQDQKRELENEE--YNRDEDEDEDEDEDEKEKEGEDEGDD--GEDEGE 448  
 QY 762 GKYQFDNLNSGNVYHFDPKSGMTQTTTDSGDDDEODADAGEEVH--VTITDHD 820  
 DB 449 GDEGEDEGE-----DEGEDEGEDEDEDEGEDEGEDEGEDEGEDE 491  
 QY 821 YQDESQSD 880  
 DB 492 EGDEGEDEGD--EGDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 550  
 QY 881 SQSD 940  
 DB 551 EDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 610  
 QY 941 SQSD 1000  
 DB 611 EDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 670  
 QY 1001 SQSD 1060  
 DB 671 GDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 725  
 QY 1061 -----DYGS---KGTL 1068  
 DB 726 VVQNPDPYTNLQKSTL 742

RESULT 13  
 ALAL\_CANAL  
 ID ALAL\_CANAL STANDARD; PRT: 1419 AA.

AC 013368;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Agglutinin-like protein ALAI precursor (Agglutinin-like adhesin).  
 GN ALAI OR ALS5.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_Taxid=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-98053977; PubMed-9393828;  
 RA Gaur N.K., Klotz S.A.;  
 RT "Expression, cloning, and characterization of a Candida albicans  
 RT gene, ALAI, that confers adherence properties upon Saccharomyces  
 RT cerevisiae for extracellular matrix proteins.";  
 RL Infect. Immun. 65:5289-5294(1997).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.  
 CC -1- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF025429; AAB88883.1;  
 DR Cell adhesion; Glycoprotein; Repeat; Signal.  
 KW SIGNAL  
 FT CHAIN 18 1419  
 FT DOMAIN 399 404 POLY-THR.  
 FT DOMAIN 408 418 POLY-THR.  
 FT DOMAIN 437 441 POLY-THR.  
 FT DOMAIN 673 676 POLY-SER.  
 FT DOMAIN 687 690 POLY-SER.  
 FT DOMAIN 700 703 POLY-SER.  
 FT DOMAIN 719 724 POLY-SER.  
 FT DOMAIN 749 752 POLY-SER.  
 FT DOMAIN 787 791 POLY-SER.  
 FT DOMAIN 869 872 POLY-SER.  
 FT DOMAIN 883 885 POLY-SER.  
 FT DOMAIN 901 911 POLY-SER.  
 FT DOMAIN 1216 1221 POLY-SER.  
 FT CARBOHYD 665 665 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 919 919 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1301 1301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1326 1326 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1419 AA: 149635 MW: 249F33F688A9D5B6 CRC64;

Query Match 5.9%; Score 331.5; DB 1; Length 1419;  
 Best Local Similarity 23.0%; Pred. No. 7.9e-06;  
 Matches 237; Conservative 116; Mismatches 366; Indels 313; Gaps 43;

QY 130 TELQKTP-----QDNHLTTEEYKSSSVSSNSIDTQAQPSHTTINRES-- 176  
 DB 66 TELNNPCVFRTASQKSVLDLADGVKATQCFYSEEFPTTSLKCTYNNNRSSIKL 125  
 QY 177 ----VQSDNVEDSHVS--DFANSKIKESNTES-----GKEENTIEOPNKVKEDSTTQPS 226  
 DB 126 GTVTLPIAFNVGCTGSSVLEDSKCFCTACTNTVFNDGSKLISI-----AVNEFKSTVDS 181  
 QY 227 GYTNIDKISNODELLNLPINYEKNARLSTSNAPSIRKRVNOLA--AQDGSVNH 284  
 DB 182 GY-----LTSREMPSLINKATLIVAPQCENGT--- 210  
 QY 285 IKVTQSITEGYDSEGCYKADAENL-----LYDTFEVDKVKSGDPTVDI 333  
 DB 211 -----SGTMGSTSTGVD--AIDCSNVHIGISKGVNDMHPPTSSEFSYTKCSCSGIS 263  
 QY 334 DKNTVPS---DLTDSFTPIPKIKDNGEIIATGTYNKKNQIYTTFTDYV--DKYENIKA 387

```

Db 264 TYQWNPAGYRPFIDAYISPS-----DNNQYOLSK-NDYTCVDYWOHAP 308
Oy 388 HKLFTYIDKSKVNN-----NKLDEYKALS-----SVKRTTVEYORENEMTAN 436
Db 309 TLKWICYKNSDAGSNQIVATRTYTDTSTANTPLPNSVKTITIELQIP--TTT 366
Oy 437 LOSMFTNIDTKNHTVEQTIYINPLRYSAKETNNVNSGNDGEGSTIIDSTIIKRYKVDN 496
Db 367 ITTSYGVWTS-----YSTKAPI-----GETATYVD----- 394
Oy 497 QNLPSNRITDYSEYEDYNDYAOAGNNDVINFGNIDSPYIIKIVISKYDKNDDYT 556
Db 395 -----VPHATTTTVESEMTGTTTTTTRNPTDSIDT-VYQVPS--PNPTTTT 441
Oy 557 --IOQVTVMOTITINE-----YTGSEFTASYDNTIAFTSSGQGGODL 596
Db 442 QFMSSEFTSTTTITLNLKGDVYAREPHNPVTTEFMSFATITETISK----- 493
Oy 597 PPKTYKIGDYWEDYDKDGIQNTDNEKPLSNLV-----TLTYPDGTSKSVRTDEG 650
Db 494 -PEGT-----DSVYAREPHNPVTTEFMSFATITET- 525
Oy 651 KYQFDGLKGLTKITFEPEGTPLKHSCTNPALD-----SEGNSVWTT----NGOD 701
Db 526 -----ITNG-----PEGTDSVYAREPHNPVTTEFMSFATITETITNKPEGTD 570
Oy 702 DMTIDSGF---YQTPKYSGLNYWYDKNKDGIDGDEKGISGVKXVTKDNGNINIST- 756
Db 571 SVYKKEPYNFTVTTEFMSFATITETITNGPEGTD-----SVYAREPHNPVTTEF 623
Oy 757 -----TTDENGKRYQFNLNSGNVYVHFDKPSGNTQTITDSCGDDGDEADGEEVHTITD 810
Db 624 WSESVATFE-----TITGGLGTD-----SIVI 646
Oy 811 HDFFSIDNGYDDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 870
Db 647 HDPLEESSSTTALES--SDSNISSAQSSESSSVESQSSIVGLSSSSDPLPSDMPSSSTG 705
Oy 871 -SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 926
Db 706 LTSESESTVSYSDSSSSTLESSTLSSDRCSSISDTTFEWD-SSSDLESTITWSSS 764
Oy 927 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 986
Db 765 IDAGSSHLVQSVKNSISTGSELSSSESESTFATDALVSDASILSDPSSVYPSSTI 824
Oy 987 SDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1038
Db 825 SSSDDPFHTIAGESSLSISFTITSTVEISSDSVSLTSDPASSFSSSSSLNDSSSSPSD 884
Oy 1039 -SDIGNSSDKST 1049
Db 885 QSDILTSSSEFT 896

```

```

RESULT 14
HYAL CANAL
ID HYAL CANAL STANDARD; PRT; 937 AA.
AC P46591;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hyphally regulated protein precursor.
GN HYAL.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RX MEDLINE=96404781; PubMed=8808922;

```

```

RA Bailey D.A., Feldmann P.J.F., Bovey M., Gow N.A.R., Brown A.J.P.;
RT "The Candida albicans HYAL gene, which is activated in response to
RT hyphal development, belongs to a gene family encoding yeast cell wall
RT proteins."
RL J. Bacteriol. 178:5353-5360(1996).
CC -1- FUNCTION: NONESSENTIAL COMPONENT OF THE HYPHAL CELL WALL
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- DEVELOPMENTAL STAGE: ABUNDANT IN HYPAE.
CC -1- INDUCTION: INDUCED SPECIFICALLY IN RESPONSE TO HYPHAL DEVELOPMENT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: Z50123; CA930485.1;
KM Cell wall; Glycoprotein; Signal; GPI-anchor; Transmembrane; Repeat.
FT SIGNAL 1..20
FT CHAIN 21..913
FT PROPEP 914..937
FT TRANSMEM 42..62
FT DOMAIN 332..576
FT DOMAIN 569..576
FT DOMAIN 577..814
FT DOMAIN 610..753
FT REPEAT 610..613
FT REPEAT 666..669
FT REPEAT 680..683
FT REPEAT 690..693
FT REPEAT 698..701
FT REPEAT 738..741
FT REPEAT 750..753
FT DOMAIN 901..904
FT LIPID 913..913
FT CARBOHYD 16..236
FT CARBOHYD 236..449
FT CARBOHYD 449..488
FT CARBOHYD 488..580
FT CARBOHYD 580..585
FT CARBOHYD 585..595
FT CARBOHYD 603..603
FT CARBOHYD 619..619
FT CARBOHYD 631..631
FT CARBOHYD 641..641
FT CARBOHYD 649..649
FT CARBOHYD 711..711
FT CARBOHYD 747..747
FT CARBOHYD 759..759
FT CARBOHYD 773..773
FT CARBOHYD 897..897
FT CARBOHYD 913..913
SQ SEQUENCE 937 AA; 93699 MW; 17802F21A1E5BA926 CRC64;

```

```

Query Match 5.6%; Score 315; DB 1; Length 937;
Best Local Similarity 21.5%; Pred. No. 2.1e-05;
Matches 189; Conservative 133; Mismatches 349; Indels 208; Gaps 30;

```

```

Oy 317 FEVDKXKSGCTPMVVDIDKNTVPSDLDSFTIIPK-----IKDNGEIIA----- 360
Db 36 FHGDVKKVSGATWAI--LGTTLCSFPGGLEVEGASLFISDNGPYLALVALVALSTLVRP 92
Oy 361 ---TGYDNKKKQIT-----YTFDYDKXENIKAHKLKLSYIDSKSVPPNNKRLD 408
Db 93 VINNGVSLNKSSTSPNSFPDGGSTFNNGEITLDDSGLYKSTAIYIARWTNG----- 148
Oy 409 VEYKTALSVNKTITVEYQRPNEKNTANLQSMFTNIDTKNHTVEQTIYIND--LRYSAK 465

```



```

Db 264 TQONPAGRRPFDATISPS-----DNNQIOLSK-KDYTCVDDYIWOHAP 308
Qy 388 HLLTSYIDSKVPPN-----MTKLDEYKALS-----SVNKKITVEYORPENRAN 436
Db 309 TLKMTGKNSDAGNSGIVATRTYTDSTAVTTLPEPNSVAKTIELQIP---TTT 366
Qy 437 LOSMFTNIDTKNHTVEQTIYINPLRSKAKETNNVNSGNGDEGSIIDSTIIVKVGDN 436
Db 367 ITTSYGVVTS-----YSTKTAPI-----GETATVIYD----- 394
Qy 497 QNPDSNRIDYSEYEDVNDYAOAGNNNDVAINFGNIDSPYIIKYSKXDPKODYT 536
Db 395 -----VPHYTTTIVSEWGTITTTTTRNPDSDIT-VVOQPS---PMPPTTTT 441
Qy 557 --IQOTVMTQTTINE-----YTGEPFASDYDNTIAFSTSGGCGDL 536
Db 442 QPMSESEFTSTTTITNLSKGDSDIVAREPHNPVTTEFEFSEFATFTIISK----- 493
Qy 597 PPKTYKIGDYWEDVDKDGIONTNDKPLSNVLV-----TLTYPDGTSKSVRDEDEG 630
Db 494 -PEGT-----DSVIVAREPHNPVTTEFEFSEFATFTET- 525
Qy 651 KYPFDLKNGLTAKITFEPEGYTPPLKHSGTNPALD-----SEGNSVWTLI---NGOD 701
Db 526 -----ITNG-----PEGTSDVIVAREPHNPVTTEFEFSEFATFTETITNKPEGTD 570
Qy 702 DMTIDSGF---YOTPKYSLGNVYADTNKDGIDGDEKIGSKVYTLKDENGNIISTT-- 756
Db 571 SYVKEPYNPTVTTEFMSSTVATFTETITNGPEGTD-----SVIAREPHNPVTTEFE 623
Qy 757 -----TTDENGKYQFDNLNGNIVHEDKPSGKTOTTTSDGDDDEODADGEVHVTITD 810
Db 624 WSESYATTE-----TITGPGKTD-----SIVI 646
Qy 811 HDPEFSDNGYDDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 870
Db 647 HDPLESSSTTALES-SDSNISSAQSSESSVEOSSSIVGLSSSDIPLSSDMPSSTG 705
Qy 871 -SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 926
Db 706 LTTSESSTVSYDSDSSSSTLSSSDRCSSISDPTTFPMS-SSSDLESTITWSSS 764
Qy 927 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 986
Db 765 IDAQSHLVQSVNSISTGSELSSSESESTFATDVLVSADASISLPTSSSYPSSTI 824
Qy 987 SDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1038
Db 825 SSSDDEPHTIAGESSLSISFTISTVEISDSVSLTSDPASPSSSSSLNDSSSSPSSD 884
Qy 1039 -SDLGNSDKST 1049
Db 885 QSDILTSSSFT 896

```

```

RESULT 14
HYRI CANAL
ID HYRI CANAL STANDARD; PRT; 937 AA.
AC P46391;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hyphally regulated protein precursor.
GN HKRI.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxId=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10261;
RX MEDLINE=96404781; PubMed=8808922;

```

```

RA Bailey D.A., Feldmann P.J.F., Bovey M., Gow N.A.R., Brown A.J.P.:
RT "The Candida albicans HYRI gene, which is activated in response to
RT hyphal development, belongs to a gene family encoding yeast cell wall
RT proteins."
RL J. Bacteriol. 178:5353-5360(1996).
CC -1- FUNCTION: NONESSENTIAL COMPONENT OF THE HYPHAL CELL WALL
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (POTENTIAL).
CC -1- DEVELOPMENTAL STAGE: ABUNDANT IN HYPAE.
CC -1- INDUCTION: INDUCED SPECIFICALLY IN RESPONSE TO HYPHAL DEVELOPMENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; Z50123; CA90485.1;
KW Cell wall; Glycoprotein; Signal; GPI-anchor; Transmembrane; Repeat.
FT SIGNAL 1..20
FT CHAIN 1..913
FT PROPEP 914..937
FT TRANSMEM 42..62
FT DOMAIN 332..576
FT DOMAIN 569..576
FT DOMAIN 577..814
FT DOMAIN 610..753
FT REPEAT 610..613
FT REPEAT 666..669
FT REPEAT 680..683
FT REPEAT 690..693
FT REPEAT 698..701
FT REPEAT 738..741
FT REPEAT 750..753
FT DOMAIN 901..904
FT LIPID 913..913
FT CARBOHYD 16..16
FT CARBOHYD 236..236
FT CARBOHYD 449..449
FT CARBOHYD 488..488
FT CARBOHYD 580..580
FT CARBOHYD 585..585
FT CARBOHYD 595..595
FT CARBOHYD 603..603
FT CARBOHYD 619..619
FT CARBOHYD 631..631
FT CARBOHYD 641..641
FT CARBOHYD 649..649
FT CARBOHYD 711..711
FT CARBOHYD 747..747
FT CARBOHYD 759..759
FT CARBOHYD 773..773
FT CARBOHYD 897..897
FT CARBOHYD 913..913
SQ SEQUENCE 937 AA; 93699 MW; 17802F212ESBA926 CRC64;

Query Match 5.64; Score 315; DB 1; Length 937;
Best Local Similarity 21.58; Pred. No. 2.1e-05;
Matches 189; Conservative 133; Mismatches 349; Indels 208; Gaps 30;

```

```

Db 149 -----LIVAVO--NOKAGNIAR-----GTAYOTITNNQICLRHOF 184
QY 466 ETNNISNG-----DE-----GSTII-----DSTIIKYVVGNO----- 497
Db 185 VPAKIKGTGVTADDEDTMIRKGLTILSEPTNHYFLKDSKSLIVHAVSSNOTFTVGF 244
QY 498 -----NLFD-DSNRIYSEYEDVTN-----DDYQALNNNNVNFNIDIPY 539
Db 245 GNGKGLITLPLTGNRHRHFRFETYPDGLDRLDALPOLYKIKGIDSKL-FRIVNSRG 303
QY 540 IIVYISKYD---PKKDDYTIQQTVMQTTINEYTGFEFTASNDYNTAFSTSSGQGD 596
Db 304 LKNAYT-YDGVVNNNEIPAVCLICTNPGSAPESDPLNTFTTSIETSSVSSATEBSV 362
QY 597 PPKTYIKGVWEDVDKDGQNTN-DNEKPLSNVLTLPDGTSKSVRTDEGCKIOFD 655
Db 363 VSESSAAV-----DSLTSLSLSKSSSDSVASTTNIESSSTIETFTMSESTD 412
QY 656 GLKNGLTATYFETPEGYPTLKHSGTNPALDSEGSNVWVINGODMTIDSGFYQPKY 715
, 413 AGSSIS-----QSESSSTAITSSETS---SSESMASSTASVTSTIETSGIYSQS-- 462
QY 716 SLGNVWYDINKDGIQGDDEKISGVKYLKDENGNIISTTTDENGKYQPDNLN----- 770
Db 463 -----ESSSNALSTEQSITSSPGSTIVASTVSTITSDENKCTEDVYTIPTV 514
QY 771 -----SGNY-----IVHFKPS-----GMTOTTT 789
Db 515 PCSTDCVPTTGDIPTMSTSYRTORTVSTITTCDEVSCSDQVYTYTNNVPHTVDAITTTT 574
QY 790 DSGDDDEDADAGEVHYITHDHFISIDNGYYDESD-SDSDSDSDSDSDSDSDSDSDSD 848
Db 575 STGDNS-----TGNSGSHHGNGNSTEGSGSGSGSGSGSGSGSGSGSGSGSG 622
QY 849 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 908
Db 623 SGSEGGSGNGSGSDSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 682
QY 909 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 968
Db 683 SEGSGSGNEGSHNGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 742
QY 969 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1028
Db 743 PGAGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 802
QY 1029 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1061
803 SGSKAGSHSGSGNEGAKTDSIGFHTESK-PGFNTGAHTD 840

```

```

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
CC TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST
CC COMPONENT OF RNA POLYMERASE II.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: THE TANDEN 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16561; CAA34560.1; -
DR PIR: S07485; RNZ02L.
DR InterPro: IPR000684; RNA_polII_repeat.
DR InterPro: IPR000722; RNA_pol_A.
DR InterPro: IPR002879; RNA_pol_A2.
DR Pfam: PF00623; RNA_pol_A; 1.
DR Pfam: PF01854; RNA_pol_A2; 1.
DR PROSITE: PS00115; RNA_POL_II_REPEAT; 9.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT ZN-FING 68 411
FT DNA_BIND 378 725
FT DOMAIN 1093 1128
FT DOMAIN 1142 1159
FT DNA_BIND 1182 1193
FT DOMAIN 1687 1694
FT DOMAIN 1258 1290
FT DOMAIN 1261 1290
FT DOMAIN 1602 1612
FT DOMAIN 1746 1759
FT DOMAIN 1806 1820
FT DOMAIN 2061 2246
FT DOMAIN 2247 2384
FT SEQUENCE 2452 AA; 278166 MW; F995E117F617A48F CRC64;
Query Match 5.58; Score 311; DB 1; Length 2452;
Best Local Similarity 19.18; Pred. No. 8.6e-05;
Matches 247; Conservative 211; Mismatches 458; Indels 374; Gaps 55;
QY 54 NSVDVKSNTDDELST-----DSNQSSDEKNDVYN 85
Db 659 NSTND-----NTDDYLEKKNAYSRSNGNHPBLSIGDINNNGVQNMSSPNNNN---N 711
QY 86 NNSGINTDNNQIIKKEETNNYD-----IEKRSDEPTESTTNDENEATFLQ--K 134
Db 712 NNNNNNNNNNN-----NNNGGINSFKFRFNMYKINLMRDSSTPSKDDNPYCSINDGK 764
QY 135 TPQDNTHLTEDEEVKSSSVSSNSI-----DPAQDSHTTINREESVQSD 181
Db 765 VIKNNELSLGIIKRTYSGSSGLIHVLMHEMGDPKTKFLSALQKVTNNMLEVGVTV 824
QY 182 NVEDSHVDFANSKIKESNTESGKEENTIEOPNKVKEDSTSQP--SGVTNIDKISNOD 239
Db 825 SCSDIASNKVLGAVRELIDSKSESVKIVE--KAKQELCEQPKSKLSYSEFETRVNNE- 881
QY 240 ELLNPLINEYNKARPLSTSAQPSI-KRYTVNOLAA--EGGS--NVNHLIKYVDOSTTE 294
Db 882 -----LNCAREMAGVASESIDERNNIFSVAGSGSIINIISQIISCVQCOQWVE 931
QY 295 G-----YD-----DSEGVKAKHDAENLI-YDVTPE-----VDDKYS 325

```

```

Db 932 GKRPFGFNHSLPHFIKFDYGPESRGFVSNSYLSGLTPQEVFFHAMGREGIIDTACKT 991
QY 326 GDTWTVD-----IDKNTPSDLTDSFTIPKIKDNSEII-----ATGTY----- 364
Db 992 SETGYIORRLIKAMEDVAVQYDFR-----VRNSYCDIIQFLYGEDMAGEYIEDQIID 1044
QY 365 ----DNK--NKQITTF-----TDYDKYENIKAHKLTSYIDSKYPPNNNTKLDVEYK 412
Db 1045 LMKLDNKINIKLYKNDEEPEFGKDYIGKNKDS--RNTTYIDYKNQNILQEFEELYK 1102
QY 413 -T-----TALSSVNTKITVEYORP-----NENRTANDOS 439
Db 1103 CKATLCKEIPPDGIRQHLPINMRILIEYAKSOFCIPFVSNNSTNNNNNNNNNNNNNN 1162
QY 440 MFTNIDTKNHTVEQTIYINPLRYSAKETNVIISGNGEGSTIIDSTIIKVKYGDNONL 499
Db 1163 SRKLMKGNLSTNNHKKRRRRRRKKNKFDKFNENNELM--SEIKKEYENNDLNM 1220
QY 500 PDSNRITDYSEYEDVT-----NDYAOLGNNNDVNIINFGNIDSPYIIVISKYDP-- 549
Db 1221 MISKG--DQSEFGKMGNEFHMGVAODMGSDLGNNNNNYN--NDDFYDDDYDD--DDYDDDD 1275
QY 550 -NRKDY--TTIQCTVMTQTTIN-----EYGEFRTASYDNT IAFS----- 586
Db 1276 YDDDDYDDDDDDDDDDSDNINIGNKRKYGTNLKNNIDENSMILNPIDVYHKVNNFEKL 1335
QY 587 -----TSSGOGGODLP-----PEKTYKIG---DYWMEVDK--- 614
Db 1336 VIIOKINSNDLVEAQNNAITILKAHRLTYLNSKLLTQTHKVSYKGLDMLQLEIEKIFY 1395
QY 615 -----DGIQNTDNKEPLSNVL-----VTLYPD-----GTSKSVRTD 647
Db 1396 KSLCHPEECYVALAOSIGEPATQMTLNTFHPAGVSKNVTGLGVRLEKELINIVKNNKTP 1455
QY 648 EDGKYQFDGLKNGLTYKITEFTPEGYPPTLKHSGNTPALDSEGNVWYTINGODMTIDS 707
Db 1456 STTYLDDMWVSNDOQKAKDILTKLEYT-TLKQLTSHAQIITDPNTTTL--EEDKSVYN 1512
QY 708 GFYQTP-----KSLGNVW--YDTNKDGIQDDEKGISGVKVTLKDENGNIISTTTDE 760
Db 1513 EYEFPPDDDTQYSLGEVNLRIQLTNH-----VNEKKLTKMLIYIIVSVSSDE 1563
QY 761 -----NGKYQFDNLNSGNYIVHFKPSCMTQTTDSCDD--DE 796
Db 1564 LDITTDNSBDLVLRIRVKYLNGEYFN-----MNYDVVDNANEQVDE 1606
QY 797 QDADGEEVHTTIDHDSISNGYDDESDS-----DSDSDSDSDSDSDSDS--SD 850
Db 1607 QEEDEEHL-----VANDRGNTEJKNSTPHPHDYNNNTNIFKSKVKNNISDINT 1657
QY 851 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS--SD 906
Db 1658 KNEBSISINSNNEQYKINSSPYNNMNNNNNNNNNNNDSSNINDIKYKNIKKEDGNEGAL 1717
QY 907 ---SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD--SDSDSDSDSD 962
Db 1718 RGGGDSNTSALFGKNKSQKEDNIVNNNNNDNDDDEEEEDFLFGDHNVSFKNTKDGKN 1777
QY 963 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1022
Db 1778 KNTNKSNNNEKNKKSNNNNNNNNNTYDDGDVNDNDNDND--DNKSDITTTIKEDNDVAFM 1836
QY 1023 SGSDSDSDSDSDSDSD--GNSSDKSTKD 1051
Db 1837 KTSTKNAEDELKKNKNIHNIHSREDED 1866

```

Search completed: March 14, 2003, 12:59:26  
Job time : 40.3513 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 41.7491 Seconds

(without alignments)  
5389,420 Million cell updates/sec

Title: US-09-147-405b-15

Perfect score: 5646  
Sequence: 1 MINKNNLTKKKPIANKSN.....FAGLGALLGRRRNRKNKN 1092

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Listing first 45 summaries

SPTREMBL.21.\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organella:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-rylous:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5646	100.0	1092	2	070022
2	4485	79.4	931	2	09K113
3	2756	48.8	1171	2	09KXW6
4	2698.5	47.8	1166	2	086489
5	2589	45.9	1141	16	099W46
6	2586	45.8	1141	16	0932F7
7	2303.5	40.8	1733	2	09K114
8	2234	39.6	1385	16	099W47
9	2093.5	37.1	953	16	099W48
10	2023.5	35.8	947	2	086487
11	1880	33.3	1315	2	086488
12	1652	29.3	989	16	099VJ4
13	1628	28.8	933	2	053653
14	1610	28.5	913	2	086476
15	1581	28.0	935	16	0932C5
16	1356	27.2	877	16	099R07

17	1471	26.1	881	2	093MH7
18	1216	21.5	1698	2	09LC00
19	785	13.9	882	16	092AK9
20	782.5	13.9	903	16	08Y697
21	671.5	11.9	487	2	09K112
22	670.5	11.9	970	11	08VBY1
23	624.5	11.1	2271	16	099QY4
24	612	10.8	2283	2	08VQ99
25	554	9.8	3394	5	077384
26	549	9.7	2570	2	092FE8
27	547	9.7	2276	2	093YV6
28	525.5	9.3	1038	16	099RD2
29	520	9.2	940	2	053682
30	502.5	8.9	961	16	099RD3
31	499	8.8	840	5	095YF6
32	498	8.8	451	3	09P785
33	471	8.3	334	5	077406
34	458.5	8.1	1946	5	097291
35	452	8.0	2910	5	026008
36	450.5	8.0	1514	5	085Y55
37	449	8.0	521	17	08T545
38	443.5	7.9	1301	5	08WSK5
39	442.5	7.8	1999	16	08YWB9
40	442	7.8	691	4	096OF7
41	441.5	7.8	1114	5	097242
42	441	7.8	2402	2	09AER7
43	439	7.8	1461	5	095259
44	435	7.7	1192	5	096127
45	430.5	7.6	4550	5	077336

## ALIGNMENTS

RESULT 1	070022	PRELIMINARY;	PRT; 1092 AA.
ID	070022		
AC	070022;		
DT	01-AUG-1998 (TREMBLrel. 07, Created)		
DI	01-AUG-1998 (TREMBLrel. 07, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Fibrinogen-binding protein precursor.		
OS	Staphylococcus epidermidis.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OX	NCBI_TaxID=1282;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HB;		
RX	MEDLINE=98261511; PubMed=9596732;		
RA	Nilsson M., Frykberg L., Floek J.I., Pei L., Lindberg M., Guss B.;		
RT	"A Fibrinogen-binding protein of Staphylococcus epidermidis.";		
RL	Infect. Immun. 66:2656-2673(1998).		
DR	EMBL: Y17116; CAA76638.1; -		
DR	InterPro: IPR001899; Gram_pos_anchor.		
DR	Pfam: PF00746; Gram_pos_anchor; 1.		
DR	TIGRFAMS: TIGR01167; LPRXG_anchor; 1.		
DR	TIGRFAMS: TIGR01168; YSIRK_signal; 1.		
DR	PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.		
KW	Signal.		
FT	SIGNAL 1 51		POTENTIAL.
FT	CHAIN 52 1092		FIBRINOGEN-BINDING PROTEIN.
FT	SEQUENCE 1092 AA; 119292 MW; 6542BC39AADB984 CRC64;		
QY	Query Match	100.0%; Score 5646; DB 2; Length 1092;	
QY	Best Local Similarity	100.0%; Pred. No. 1.5e-182;	
QY	Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
DB	1 MINKNNLTKKKPIANKSKVAIRKTYGTASIVGATLLFGIGHNEAKAEENSVDYK 60		
QY	1 MINKNNLTKKKPIANKSKVAIRKTYGTASIVGATLLFGIGHNEAKAEENSVDYK 60		
QY	61 DSNFTDELSDSDSSDEEKNDYINNNOGINTDNNQIIRKEETNITGIEKSESDRTES 120		





Db	396	KTYP-NETSUNLFEAFAAGKETSONVYVDVQDPKAVHSDNSIQSIFTKLDEDKQTEIQOLIV	454
QY	458	NPLRYSAKETNNVNSGNC-----DEGSIIDDSIIKKYKVGDNQNDPSNRIDY	508
Db	455	NPLKKSATNTKVDIAGSOVDYGNIKLKGNSITIIDNTEIKKIVKVSQDLQPSNRIDE	514
QY	509	SEYEDVTND-DYAOQLGNNDVNIINFGNIDSPYIIKYSIKYDPKDDYTTIQOTVMTQTTI	567
Db	515	SOYEDVTISQDNNKSFSENNNATLDFDINSAYIIKIVKSYTPTSDELDIAGTSMRT-	573
QY	568	NEYGEERTASVONTIAFSRSGGGGGD--PREPKYIKIDYVWDVDRKQIONTDNNEP	626
Db	574	DKY-GYINYGISNFPIVTSMDTGGGGYKPEEKLKIDGYWEDVKCGVGTOSKEP	632
QY	627	LSNVLVLTYPDGTGSKSVRTDEGKYQDFGLKNGLTYYKITTETPEGYPTTLHSGTNPAL	686
Db	633	MANVLVLTYPDGTGSKSVRTDANGHIEFGSLKGEYTYKFEPIPGYLTFTKXNGTTDGGK	692
QY	687	DSEGNVWVTTINGODMTIDSGFYQTPPKYSLGNVYVDTNNKDGIOGDEKIGSKAVYTLK	746
Db	693	DSNSSVTVKINGKDDMSLDTGFEKPKYMLGDIYWEDTKDGIOANEPGIKDVYTLK	752
QY	747	DENENIISTT-----	757
Db	753	DSFGKVIQTTTDSAGKYEFTDLONGNVYEFETPAGYTPYKNTTADDKNSGLTTTGV	812
QY	758	-----	757
Db	813	IKDADNMTLDRGFYKTPKYSLGDIYVYDSNKKDKOSTEKGIKDVTVTLQANKGEVITTT	872
QY	758	-TDENGKYQFDNLNSGNYIVHFDKPSGMTQTTTDSGDDDEQADAGEVHYTPTDHDPSI	816
Db	873	KTDENGKYRFEDNLDSGKYKVFEEKPAGLQGTVNTTEDB-KADGGEVDTVTTDHDFTL	931
QY	817	DNGYDDESD	876
Db	932	DNGFFEDT-----SD	980
QY	877	SD	936
Db	981	SD	1044
QY	937	SD	996
Db	1041	SD	1100
QY	997	SD	1055
Db	1101	SDSDSDGKRTPPVKPMS-----TTKDHNNKAKALPE	1133
QY	1056	TGANEDYSGKGTLLGTLFAGIAGALL-GKKRKNRK	1089
Db	1132	TGSENNGSNNATLFGGLFALGSLLLFGRRKKQNK	1166
RESULT 5			
Q99M46	ID	PRELIMINARY;	PRT; 1141 AA.
Q99M46	OS	PRELIMINARY;	PRT; 1141 AA.
AC	Q99M46		
DT	01-JUN-2001 (TEMBLrel. 17, Created)		
DT	01-JUN-2001 (TEMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (TEMBLrel. 21, Last annotation update)		
DE	Seq-asp rich fibrinogen-binding, bone stalooprotein-binding		
DE	protein.		
GN	SDRE OR SA0521.		
OS	Staphylococcus aureus (Strain N315).		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Staphylococcus		
OX	NCBI_TaxID=158879;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
FX	MEDLINE=21311952; PubMed=11418146;		
FA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,		





[illegible]

```

QY 324 KSGDTMTVDIDKNTVPDLTDSFTIPKIKDNCSGELIATGTDNKKNQITTFEDYDKKE 383
Db 412 KSGDTFTIKYGOYIRPGGILEPAIKTOLKSKDGSIVANGYDKTMTTFTTFTTNYDOXQ 471
QY 384 NIKAHKLITLSYIDKSKVPNNNTKLDVEYKTALSVNKTITVEYORPNENRANLQSMFTN 443
Db 472 NITGSEFDLATPKRETAIKONQNYPMIEVTAINEVYKKDFIVDGKKNKMTT---AAVAN 528
QY 444 IDTKHHTVEQTIYINPLR-----YSAKETN----- 468
Db 539 YDNVNNKHNEVYLLNONNPNRYAKYFSTVKNGEFIPGEVKYEVTDITAMVDSFNPDLN 588
QY 469 -----VNISGND-----EG 478
Db 589 SSVNKKVTSQFAPKVSADGTRVDINFAKSMANGKATITVQANRPITGTGVNTEYELTRPG 648
QY 479 STIIDD-----STII-----KVKVGD-----NQ----- 498
Db 649 TTNNTDFYGTSTYTYLNGSSTAGDNPTSLSDYVWLKKNKNGVODDEKGLAGVYV 708
QY 499 -LPDSN-----RI-----YDYSEEDVT-----NDVYQLGNNNDVN----- 529
Db 709 TLKDSNNRELQRTVTDQSGHYQFDNLONGTYVEFAIPNNYPSPANNSTNDAIDSGER 768
QY 530 -----INFGNIDSPYI---IKVISKY-----DPNKD-----DVTTLQ 558
Db 769 DGTAKVYVAKGTIN--NADNMVYDGTGYLTPKYNVGDYWEIDNKGIDDDNEKGISGVK 826
QY 559 OTV-----TMOTLINEYGEFRTASYN--TIARST-----SSGOGOG-LDPRE 599
Db 827 VILKKNNGTIGITTTDSNGKIEFTGLENGDYITEFEETGEGYTPPTQNSGSDGKSNGT 866
QY 600 K-----TYKIDGYWEDYDKDGIQNTFNDEKPLSNVLTLYTPDG 639
Db 887 KTVTVKADANKTIDSGFAPKPTYNLGDYWEIDNKGIDQ--DSEKIGISGVKTYLKDKN 944
QY 640 TS-KSVRTBEDGKYQFDGKLNGLTYKITEFTPGYTPPLTKHSGTNALDSEGSVWVTIN 638
Db 945 NALGTTTASGHYQFKGLENG--SYVEFEETPGYTPPTKANSQODTIVDSNGITTTGILN 1003
QY 699 GODDMTIDSGFYOTPKYSLGNYWYDTPNKGIDGDEKGISGVKTYLTKDNGNIISTTTT 758
Db 1004 GADNLTIDSGFYTPKYSVGDYWEIDNKGIDDDNEKGISGVKTYLTKDKNIISTTTT 1063
QY 759 DENGKYOFPNLNSGNYIVHDKPSGMTQTTTDSGDDEQDADGEVHVITTDHDDPSIDN 818
Db 1064 DENGKYOFPNLDSGNYIIHFEKPEGMTQTTANSNGNDEKADGEDVAVITTDHDDPSIDN 1123
QY 819 GYVDD----- 823
Db 1124 GYFDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1183
QY 824 ----- 823
Db 1184 SDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1243
QY 824 ----- 823
Db 1244 SDADSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1303
QY 824 ----- 823
Db 1304 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1363
QY 824 ----- 823
Db 1364 ADSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1423
QY 824 -----ESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 842
Db 1424 SDSDSDSDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1483
QY 843 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 902

```

```

Db 1484 SDSDSDADSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1543
QY 903 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 962
Db 1544 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1603
QY 963 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1022
Db 1604 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1663
QY 1023 SGSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1082
Db 1664 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1723
QY 1083 KRRKRNK 1091
Db 1724 RRRK-KDNK 1731

RESULT 8
Q99W47 PRELIMINARY: PRT: 1385 AA.
AC Q99W47:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDRD OR SA05520.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria: Firmicutes: Bacillus/Clostridium group: Bacillales:
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ul Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshida K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003359; BAB56724.1;
DR EMBL: AP003351; BAB41751.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRfams: TIGR01167; LPTXg_anchor; 1.
DR TIGRfams: TIGR01168; YSRK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 1385 AA; 149643 MW; 2A0CD27733B3C1D CRC64;

Query Match 39.6%; Score 2234; DB 16; Length 1385;
Best local similarity 38.5%; Pred. No. 7,6e-68;
Matches 565; Conservative 166; Mismatches 277; Indels 458; Gaps 46;

QY 1 MINKKN-NLTKKKPIANKSNKYAIRFTVGTASIVGATLLEGLGHNKAEKNSVDYV 59
Db 1 MLNRENKTAIRKGMVSNRLNKESIRKRYTGATSIILVGTILFGLGQKAEKSTKEL 60
QY 60 KDSNTDELSDSNQSSDEKNDVYNNQSIINDNNQIILKEETNNYDGIKRSDEPTE 119
Db 61 NEATF---SASDQSSD--KVDMOQLNOEDNTRKNDQ-----KENVSQGN 102
QY 120 STTNVDEN-EATFLQKTPQDNTLTHEEVKSSSVSSNSIDTAQOPSHTTINREESVQ 178
Db 103 TTSGNKSIRKESVQSTTGKAKVEYSTAKSDQASPKSTNEDLNKQ-----TISNGGLQ 157

```

```

QY 179 TSDNVEDSHVDFANSKIKESNTESGKEENTIEQPN-KVKEDSTTSQPSGYTNIDEKISN 237
D 158 -PDLEENKSVVN-----VQPTNEENKRVDAKTESFTTLNVKSDAIKS-----NAETLVDN 205
QY 238 ODELLNLPINEYENKARPLSTTSAPPSIKRVTVNOIAEQQS-----NNHLIKYTDQST 293
D 206 NSNSNN-----ENNADILPKSTAPKSLNTRMAAIOQPSNDSKANNVNDLSTNTTLTV 259
QY 294 EGYDDSEGVAKAHAEENLIYDTPFEVDKVKSGDMTMDIDKNRY-----PSDLTDSFT 347
D 260 VDAHQNSKTIIVPAODYLSKQIT--VDDKVKSGDFTIKY--SDYQVYGLNEDIKN-- 313
QY 348 IPKIKD-NSCEIIATGYDNKNKOITTYTDYVCKENIKAKHLKLTYSIDSKVYNNNTK 406
D 314 IGDIKDPNNGETIAFAKHDANNLITFTFYDVRFNKVKGINSTMDADTIPVD--K 371
QY 407 LDVEYKALSSVKNKTIIVEYQRP--NENRTANLOSMTNIDTKNHT-----VEQTIY 456
D 372 KDVEFSVTIGQITTTTADITYPAYKEADNNSIGSAFT--ETVSHGVNVEDPGYVQVYV 429
QY 457 INPLKYSAKETNNVNSGNGEGST-----IIDSTIIKYYKKGDNQNLDSRIYD--SE 510
D 430 VNPMDKLGAKLKEAYHPRPTNIGQINQNTNIKIYRVP3GYTL--KKGVDVNTND 486
QY 511 YEDVTND--DYAOLGNNDVNIENFGNIDSPYIIKIVISKYDNKDDYTTIOQTVMQTNIN 568
D 487 LVDVTDEKKNMKTGNSQSVNLFEGDITSAYVWNTKFOYTNSSEPLVYQMATLSSIGN 546
QY 569 EYTGFEFRASIDNTIAFSTSGGOGGDLPEKTYKIGDYWEDVDKDIQ----- 618
D 547 -----KSVSTGNMAGFTNNQSGAG--OEVYKIGYVWEDTKNGVQELGEKGVNV 596
QY 619 -----NTN----- 621
D 537 TVTVFDNNTNKKVGEAVTKEDGSYILPNLPNGDYRVEFSNLPKGYEVTSPKQGNNEELDS 656
QY 622 -----DN-----EXPLSNVLT----- 634
D 657 NGLSSVITVNGKDNLSADLGIIKPKYNLGDYWEDTKNGIDODDEKISGYTVTLKDN 716
QY 635 -----TYPDG----- 639
D 717 GNVLTVTTDADGKXKFTDLNNGNKYVEFTPEGTYPTTVTSGSDIEKDSNGLTTGVIN 776
QY 640 -----TSSV-----RT 646
D 777 GADNMTLDSGFYKTPKYNLGNVWEDTKKDGKODSTEGISGYTVTLKNGEVLQTKT 836
QY 647 DEDGKYQFDGLKNGLTLYKITEFTPEGYPTLKHSQTNALDSEGS----- 692
D 837 DKDKGYQGTGLENG--TYKVEFETPSGYPTVOGSGTIDSGIDSNGSTTGIVIKDKNDITD 895
QY 633 -----VW-----VTI-----NSQ----- 700
D 896 SGFYKPYNLGDYWEDTKNKGNOVDKDEKISGYTVTLKDEHDKLKTVTTDENKGYOFT 955
QY 701 -----DDMTIDSGFYQTPKXS 716
D 956 DLNNGTYVEFETPSGYPTSVTSNDTEKDSNGLTTGVIKDANMTLDSGFYTPKXS 1015
QY 717 LGNVWYDNTKDGLOGDEKISGYKVTLKDENGNIISTTTDENGKYQFDLNSGNIV 776
D 1016 LGDYWYDSNKGDKODSTREKIKDVKVILLNEKGEVITTKTKDENGKXRFDLDSGKXKV 1075
QY 777 HFDKSGMTQTTSDGDDDDADAGEVAVITTDHDFSIDNGYVDE--SDSDSDSDSDS 835
D 1076 IFEXPTGLTGNTTEED--KDADGCEVAVITTDHDFLLDNGYEEETSDSDSDSDSDS 1134
QY 836 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 895
D 1135 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1194

```

```

QY 896 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 955
D 1195 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1254
QY 956 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1015
D 1255 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1309
QY 1016 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1064
D 1310 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1359
QY 1065 KGTLLGLTLAGLALL--GKRRKNRK 1089
D 1360 NATLFGGLFALAGLTLFGRKKQK 1385
RESULT 9
QY 099M48 PRELIMINARY; PRT; 953 AA.
AC 099M48:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDRG OR SAV0561 OR SA0519.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Itoh T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murekami H., Hoshiyama A.,
RA Mitutani U.I., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hiramata H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hirayasu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus ";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003359; BAB56723.1;
DR EMBL; AP003131; BAB41750.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 953 AA; 103292 MW; 729A7169A074A1E5 CRC64;
Query Match 37.1%; Score 2093.5; DB 16; Length 953;
Best Local Similarity 43.8%; Pred. No. 2.8e-63;
Matches 488; Conservative 140; Mismatches 296; Indels 189; Gaps 26;
QY 2 INKNNLLTKRRKPIANKSNKAYAIRFTVGTASIVIGATLLFGLGHNKAEENSQVDYKD 61
D 1 MNKKATATNRKGMNRLNRSKIRYSVGTASILVGTLLFGLSHEKAAE----- 52
QY 62 SNTDELSDQSDDEKNDVYNNQSIINDNNOIITKEETNNYDGIKRSDEPTEST 121
D 53 -HTNGLNQSNETTAPSEN----- 74
QY 122 TNVDENEATFLQKTPQDPTHLTEEEVKSSSVESNSSIDTAQOPSHTTINREESVQSD 181
D 75 EKVDNRQL-----KDNVQTA-----TADQPKVT----- 97
QY 182 NVEDSHVDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEK---ISNQ 238

```

```

Db 98 -----MSDSATVKTSSMMS-----PQNAASOSTQTSNVTNDKSSSTYSNE 142
QY 239 DELNLPIREYEN-KARPLSTTSAOPSIRKRVYNOLA-EGSNVNLKIVYDOST-----292
Db 143 TDKSNL--TQAKNVSTPTTITIKORALRMNAVNTYAAQOGTINVNDKHFNTIDAIOK 200
QY 293 -----TEGYDDESEGVAKAHADENLIYDVFEDDKVKSQDPTMTVIDIKNTVPSDLT 343
Db 201 GHVNTKTGTEFMATSSDVLK-----LKANYTIDDSVKEGDEFTFPGYGFRRGSVR 252
QY 344 DPTTPKIDNSGEIATITGYNKKQITTYTDDYDKENIKAHKLKLSYIDKSKVPPNN 403
Db 253 LPSQTONLNAOGNIIAKIYDSKNTTITTYFTNYVDQYTNVSGSFEQVAFAKRENATD 312
QY 404 NPKLDVEYTAALSSVKKITTEVEORPENRNTANLQSMFNIDTKNHTVEQITIN-PLRY 462
Db 313 KTAIVKEVTLGNDYTSKDYIVDI---GNQGOOLISSTYINNEEDSRMNTVYVNPCKT 369
QY 463 SAKETNV-NISGNGDEGSTIIDSTIIKVKYKVDNQNLPDSHRIYDYSEYEDVTD-DYA 520
Db 370 YTKETFTVTLT-----GYKFNPDAKNFKEYETDQNOFVDS-FTPTPSKLKDYTGQFDVI 423
QY 521 QLGNNNDVINFGN---IDSPYIIVISKYDPNKDVTYITQOTVTMOTTINEYTGEEFT 576
Db 424 YSNDNKTATVDLILNGOSSSDKOYIIQOVAVPDSSTDNCKIDYTLTETONKGSMSN----479
QY 577 ASYDNTIATSTSSGQGGDLPEPKYKIDGYWEDYDKGIONTNDNEKPLSVLTLTY 636
Db 480 -SYSNNGSSTANGD-----QKYNLGDYWEDTNKQKDA--NEGKIGCVYILKD 529
QY 637 PDGTS-KSVRTDEDEKYYQDGLKNGLYKITEFTEPEGYPTLKHSGTNPALDSEGSVWY 695
Db 530 SNGKELDRTTIDENKGYQTGLSNG-TYSVEFSTPAGYPTTANAGDAVDSGLTTTG 588
QY 696 TINGQDDMTIDSGFYQTPKYSILGNIYWTNKGIGQDDEK3ISGYKVTILKDENGNIST 755
Db 589 YIKDADNMTLDSGFYKTPKYSILGDIYVWYSNKGKQDSTEGIKGYKVTILQNEKEGYGT 648
QY 756 TTTDENKGTQFONLNGNIVHFDPKPSGMTQTTTDSGDDEODADGEVAVTTTDDHDFS 815
Db 649 TETDENKGTAFDNLDSGKRYKITEKPAGLTQGTNTTEDD-KPADGGEVAVTTTDDHDF 707
QY 816 IDNGYVDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 875
Db 708 LDNGYEEFT-SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 766
QY 876 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 935
Db 767 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 826
QY 936 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 995
Db 827 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 886
QY 996 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1055
Db 887 DSDSDSDAKHTPTKPKSTV-----KDHKTAKALPE 918
QY 1056 TGANEDYGSKGTLLGTFLAGLALL-GKRRKRRK 1089
Db 919 TGSENNNSNNGTLFGGLFALGLLFGRRKKONK 953

```

```

RESULT 10
ID 086487 PRELIMINARY; PRT; 947 AA.
AC 086487;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE SRC protein.
GN SDRC.

```

```

OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxId=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN.
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCreia K., Ni Eldhin D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RT "three new members of the serine-aspartate repeat protein multigene
family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL: AJ005645; CAA0650.1; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor.1.
DR TIGRFAMs: TIGR01167; LPTG_anchor.1.
DR TIGRFAMs: TIGR01168; YSTRK_signal.1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 947 AA, 10288 MW, 366FD6E35121534 CRC64;

Query Match 35.8%; Score 2023.5; DB 2; Length 947;
Best Local Similarity 42.3%; Pred. No. 6.2e-61;
Matches 470; Conservative 158; Mismatches 297; Indels 185; Gaps 28;

QY 2 INKNNLTKRKPIANKSKRYAKRFTVGTASIVIGATLLFGLHNEAKAENSVDVXD 61
Db 1 MNKKATATNRKGMIPRLNKFISIRKYSVGTASIVGTTLIFGLSGHEAKAAE-----52
QY 62 SNTDDELSDNSDSSDEKNDVNNNNOINTDNNQITKEEFNNTDGLKEKREDTEST 121
Db 53 -HTNGILNOSKNTTAPSEKTKKVDYRQLKDNTO-----T 88
QY 122 TVNDENAEPLQKTPDQNTHLTEEVKSSSVSSNSIDTAQPSHTTINREESVQSD 181
Db 89 ATADQPKVT-----MSDSATVKTSSNM---QSPQNAATNOSTTTSNV 129
QY 182 NVEDSHVDFANSKIKESNTSEKREENTJEPQPKVKEDSTSQPSGTINIDEKISMODEL 241
Db 130 TTNKDSSTYSNTEKDSNLTQA-----KDVSTT-----157
QY 242 LNLPIREYENKARPLSTTSAOPSIRKRVYNOLA-EGSNVNLKIVY--DOSITEGY-D 297
Db 158 -----PKTTIIPRTLNRMAVNTYAAPQOGTINVNDKHFNTIDAIOKHVN 204
QY 298 DSEGVAK-AHADENLIVDYFEVDKVKSGDPTMTVIDIKNTVPSDLTDSFTIPKIDNS 355
Db 205 QTGKTEFMATSSDVLKLNANYTIDDSVKEGDEFTFPGYGFRRGSVRLPSQTONLYNAQ 264
QY 356 GEIITAGTYNNKKQITTYTDDYDKENIKAHKLKLSYIDKSKVPPNNTKLDVEYKTL 415
Db 265 GNIIAKGIYSTNTTITTYFTNYVDQYTNVSGSFEQVAFAKRNATTDKAYKKEVTLGN 324
QY 416 SSVNKTITVEYQRPENRNTANLQSMFNIDTKNHTVEQITIN-PLRYSAKETNV-NISG 473
Db 325 DTYSEETIYDI---GNKKAQPLISSTYINNEEDSRMNTVYVNPCKATYTKQIFVTULT- 380
QY 474 NGDEGSTIIDSTIIKVKYKVDNQNLPDSHRIYDYSEYEDVTDYALQGNNDV 528
Db 381 -----GYKFNPDAKNFKEYETDQNOFVDSFTPTPSKLKDYTGQFDVI---YSNDNKTATV 433
QY 529 NINEGNIDS--PYIIKYI-----SKYPNKDVTYITQOTVTMOTTINEYTGEEFTASYDN 581
Db 434 DLKGGTSSNKQYIIQOVAVPDSSTDNCKIDY-----TLDDTKRYSM---SNSYSN 483
QY 582 TIAFTSSGQGGDLPEPKYKIDGYWEDYDKGIONTNDNEKPLSVLTLTYPGTGS 641
Db 484 VNGSSTANGD-----QKYNLGDYWEDTNKQKDA--NEGKIGCVYILKDSNGKE 534
QY 642 -KSVRTDEDEKYYQDGLKNGLYKITEFTEPEGYPTLKHSGTNPALDSEGSVWVYTINGQ 700
Db 535 LDRITTDENKGYQTGLSNG-TYSVEFSTPAGYPTTANAGTDAVDSGLTTGVYIKDA 593

```



```

QY 979 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1038
D 1194 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1253
QY 1039 SDGNSD-----KSTKD-----KLPTGANEDEGS-KGTLGLTFLAGLALLL-GKRRK 1086
D 1254 SDACKHTPKPMSTTKDHNNKAKALPEFG-NENSGSNNAFLFGGLFALGSLLEGRKK 1312
QY 1087 NRK 1089
D 1313 PNK 1315

RESULT 12
Q99VJ4 PRELIMINARY; PRT; 989 AA.
ID 099VJ4:
AC 099VJ4:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE Fibrinogen-binding protein A, clumping factor.
GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; Pubmed=11418146;
RA Kirada M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsunaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani O.Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kato C., Sekimizu K.,
RA Hirakawa H., Kubara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003131; BAB41975.1; -.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMs: TIGR01167; LpxTG_anchor.1.
DR PROSITE: PS00402; BPD_TRANS_P. INN_MEMBER: UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 989 AA; 102407 MW; DA6E807539623467 CRC64;

Query Match 29.3%; Score 1652; DB 16; Length 989;
Best Local Similarity 37.9%; Pred. No. 1.9e+48;
Matches 432; Conservative 155; Mismatches 330; Indels 224; Gaps 36;

```

```

D 247 TTVPHQAGYKLLVGFSSVPSAAVKGDTFKITVREKLNKGVSTAKVPPIMAGD-OYLA 305
QY 361 TGYDNKNKQITFTFTYVDKYEENIKAKLKTSTYIDSKVNNNTKL-DVEYKTL--SS 417
D 306 NGVIDS-DGNVITFTYVDKYEENITANTIPATID-----PENVTKGNAVLTIGIGINT 360
QY 418 VNKITVEYORPNENRANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNV-----N 470
D 361 ASKTVLIDYEKYGQFHNLSIKGTIDQIDKTNNTFYQTIVNP-----SGDNVVLPAALGN 415
QY 471 ISGNDEGSTIIDSTIIKYKVGDNONLPSNKRITDYSEEDYTDNDYADLGNNNDVNI 530
D 416 LIPNTKSNALIDANKTDLKIVRV-DNANDLESYVNPSEDEYTNQVRISFPANQKV 474
QY 531 NF-----GNIDSPYIIKIVSKYDPNKKDVTYTIQGVYTMQTTINETGEF--RTASVDNTIA 584
D 475 EFPIDDDQITTPYIVVNGHIDPASTG-----DLALRSTFYGDSNFTMRSMNDNEVA 528
QY 585 FSTSSGOGG-DLP--PEKTYKIDY--VWEDVYKDGIQNTNDEKPLSNVLTLYTPDG 639
D 529 FNGSGSGDGIDKPVVPEQPEGPRIPIRPSDS-----PGS 567
QY 640 TSKVYRDEDECKYQFDGLKNGLYKITFETPEGTPTLKHSCTNPALDSEGSVWYTING 699
D 568 DSGDSNSDSGS-----DSGSDSTSDSGSDASDSAS----- 601
QY 700 QDDWTIDSGFYQTPKYSLGNYVWYDTNKGDIQGDDEKISGVKTKLKNENGIISTTTD 759
D 602 DSDASDS-----DSASDSASASDSAS-----DSASASDSASD 638
QY 760 ENGRYQPDNLNSGNVYHFDKPSGMTQTTTDSGDDDEQADAGEVHTITDHDDEFIDNG 819
D 639 SDSASDSASASDSASDSASD--SDGASDSDSDS-----SDSDSDSDSD 689
QY 820 YYDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 879
D 690 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 749
QY 880 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 939
D 750 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 809
QY 940 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 999
D 810 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 869
QY 1000 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1046
D 870 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 929
QY 1047 -----KSTKDKLPTGANEDYGSKTLTGLTFLAGLALLLGRKKRKN 1090
D 930 PNSPKNGTNASKNKAKEKLEPLPTGSEDE--ANTSLIWGLLASLGLLFRKKRKN 987
QY 1091 K 1091
D 988 K 988

RESULT 13
Q53653 PRELIMINARY; PRT; 933 AA.
ID 053653:
AC 053653:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]

```

```

SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RX MEDLINE-94224142; PubMed-8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor(fibrinogen receptor
of Staphylococcus aureus."
RL EMBL: Microbiol. 11:237-248(1994).
DR MBL: 218853; CAAT9304.1;
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRfams: TIGR01167; LpxG_anchor; 1.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 933 AA; 97058 MW; EB51A6DE2FF759F4 CRC64;

Query Match      28.8%; Score 1620; DB 2; Length 933;
Best Local Similarity 37.6%; Pred. No. 1.2e-47;
Matches 420; Conservative 153; Mismatches 313; Indels 230; Gaps 35;

18 KSNKYAIRKFTVGTASTVIGATILFG-LGHNEAKAEENSQVYKDSNTDDELSDSDSQS 76
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   5 KKEHAIIRKKSIGVASYVGLTIGFGLSSKEADASENSV-----TQSDSASN 52

QY 77 DEEKNDYINNNSQINTDNNQIIKKEETNNYDGEKRESDTESTVNDENAEFLQKTP 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 53 ESKNDSSVSAAKTDITDV-----SDTKTSNTN--NGETVAQNP 93

QY 137 QDNTHLLEEVEKSSSVESNSSIDTAQOPSHHTINREESVCTSDNEDSHVSPFANSKI 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 -----AQOETQSSSTNATEETEPVTEGATTTTNCANPATQOS-SNNABELVN--- 143

QY 197 KESNTEGKRENTTEQNNKKEKEDSTQSPSGYTNIDEKISNDLELN-LPINYEKARP 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 QTSSETTFNTNTV-----SSVNSPQNSTAENVSTQDSTREAPSN--NESAP 191

QY 256 LSTSAOPSIKRYVN-----QLAEGGSNVNHLIKYTDQ--SITGYDSEGVK 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 QSTDAKNDVYNAQVNSARBRMRAFLAAYAAADPAAGTDITNTLTNTVGTI-DSGTIV 250

QY 305 AHAENLIYDVEFEVDKVGSGDTMTVDIDKNTVPSDLTSTFIPIKIKNSGELIATGY 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 PHQAGYVKLWYGFSPNSAVAKDFPKITVPKELNGLVSTAKVPIMAGD-QVLANGVI 309

QY 365 DNKKQITVETFDVVDKYEENIKAHKLTSYIDSKVNNNTKLV-DVEKTAALS--VNKT 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 DS-DGNVITFTDVNKKDKVATLIPATID---PENVKKGNVLAIGIGSTTANKT 364

QY 422 ITVEYQRPENRTANLQSMFTNIDTKNHTVEQTIYINPL--RYSAKETNNISGNGDEGS 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 VLVDYEKGFYNSIGKIDQIDKTNNTYKQTIYVPSGSDNVIAPVLTGKIKNTSMA 424

QY 480 TIIDSTIIVYKVGNDQNLPSDSRIYDYSEYEDVTNDYQAQIGNNNDVINFG---NI 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 LIDQNTSIRIKYKVNAAADISESYFV-NPEFEDVTNSVNTFPNPQYKEFTPDQI 483

QY 536 DSPYIIVISKYDQPN-KDDYTTTIOQTYMOTINETGE--FRASADNTIAPSTSSGOG 592
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 484 TTPPIVAVNGHIDPNSKGD-----LALRSTLYGNSNIIWRSMSDNEVAFAFNGSGG 536

QY 593 QGDLPKRTYKIGDYVEDYKDIQNTNDEKPLSNVLVLTLPDQTSKSVRTDEGKY 652
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 537 DG-----IDKPYVP-----EQP-----DEGEI 554

QY 653 QFDGLAKGLTKYTEETPEGYPTLKHSGTNPALDSEGSNVWVTINGQDWTIDSGEYQT 712
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 E-----PIPEDSDSDPGSDSGSDS-----NSDSDSGSDSGSDST 586

QY 713 PKYSLGVVWYDTNKGDIQGDDEKIGSVKVTLKDENGNIISTTTDENKYPDNLNSG 772
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 587 S-----DSGSDSASDSD-----SASDSDSASDSDSASDSDS 615

QY 773 NYIVHFDKPSGMDQTTTDSGDDDDQADGGEVHTITDHDDFSIDNGYVDESDSDSDSD 832
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 616 D-----SASDSDSDNDSDSDSDSD-----SDSDSDSDSDSDSDSDSDSDSDSDSDSD 658

QY 833 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 892
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 659 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 718

QY 893 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 952
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 719 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 778

QY 953 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1012
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 779 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 838

QY 1013 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 1055
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 839 SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 898

QY 1056 TGANEDYGSKGTLLGTFLAGLGAALLKRRKRNK 1091
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 899 TGSEDE--ANTSLIWGLASIGLLFRKKENKDK 932

RESULT 14
ID 086476 PRELIMINARY; PRT; 913 AA.
AC 086476;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Clumping factor B precursor.
GN ClfB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_Taxid=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RA Foster T.J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RA N. Eldhun D., Perkins S., Francois P., Vaudaux P.;
RT "Clumping factor B(ClfB), a new surface-located fibrinogen-binding
adhesin of Staphylococcus aureus."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
EMBL: AJ224764; CAAL2115.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRfams: TIGR01167; LpxG_anchor; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 44 POTENTIAL.
FT CHAIN 45 913 CLUMPING FACTOR B.
SQ SEQUENCE 913 AA; 97247 MW; 38077C321F8D3B61 CRC64;

Query Match      28.5%; Score 1610; DB 2; Length 913;
Best Local Similarity 36.9%; Pred. No. 4.6e-47;
Matches 426; Conservative 126; Mismatches 277; Indels 324; Gaps 36;

15 IANKSNKYAIRKFTVGTASTVIGATILFG-LGHNEAKAEENSVDYKDSNTDDELSDSDSQ 74
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 LSNKKKKYSIRRTTGTIVIGATILFGIGNHQAQASEQS-----NDTQSSKNN 59

QY 75 SSDEKNDVYNNQSTIT--DDNNQIIKKEETNNYDGEKRESDTESTVNDENAEFL 132
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 SADSEKNNMIETPO-LWTTNDNTSDISANTNSANVDSITKPMSTQTSNTTT--TEPAST 115

QY 133 OKTPQD---NTHLLEEVEKSSSVESNSSIDTAQOPSHHTINREESVOTSDNVEDSHVS 189
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
DB 116 NETPOTAIKNGATAKMODOTVPOEGNSQVD-----NKTNDANSTIATNSELKNSQTL 169
QY 190 DFANSKIKESNTESGKEENTIEOPNKKVREDSTTSQPSGYTNIDEKISNODELLNPINEY 249
DB 170 DLPGS-----SPQTSIAQGSTSKPSVTRAVRSLAVALPEYVN----- 206
QY 250 ENKARPLSTSAQPSIKRVTYNQJLAEOGSNNHLIKTDSITIEGYDSEGVIAHDAE 309
DB 207 -----AADAKGTNNV-----DKVTASNFLEKTEFTPDPOSG 237
QY 310 NLIVDFEVDOKVSGDPMY-----DID-----KNVPSGLDSTFIPIKONS 355
DB 238 NTFMANFTYVDKVSQDFTAKPLDPSLTGNGVDVYSSNNMTP-----IADIKSTN 289
QY 356 GEIATGYDNKNKOITFTFDYDKYENIKAKHLKLTYSIDSKVPNNNTKLDVEKTL 415
DB 290 GDVAKATYDILTKTYTFTDYVANNKENINGQFSLPFTDAKAPKSGT-YDANININAD 348
QY 416 SSVNKTIVEYORP-----NENFTANLQSMFTNIDTKN-HIVEQTIYINPLRYSAKETNV 469
DB 349 EMFNKITYNASSPIAGIDKPNGANISSQILIGVDTASGQNTYKQTVFVAPKORVLGNTVY 408
QY 470 NISGNGDEGSTIIDSS-----TLIKYKYGDNQNLPS-----N 503
DB 409 YIKGYQDK-----IESSGKVSATDTKRLRFEVNDTSKLSDSYADPNDSNLKEVTDQEN 464
QY 504 RIYDSEYEDYTNDDYAOGLGNNDVNNINGNIDSPYIIKIVISKYDPNKKDYTT--IQQTV 561
DB 465 RIY-----YE-----HPNVASIKRGDILTKTYVVLVECHYDGTGKLNKTOYQENY 509
QY 552 TMOITINEYTGEPFASIDNTIAFTSSGQGGDLPEKTYTIGDYWEDVDKQIQNTN 621
DB 510 -----DEVNDRYCI--FGW-----NN 524
QY 622 DNEKLSNVLTLYTPDGTSKSVRTDEDEKGYOPDGLKNGLTIKIFEPREGYPTLKHSG 681
DB 535 EN-----VVRYGGA-----DG-----DSAVN-----PKDPTP----- 548
QY 682 TNPALDSEGNISYVWYINGODDWTIDSGFYQPKYSIGYVWYDTNKGIGQGDDEKISGV 741
DB 549 -GPPYDPEPSP-----DPEPEPTPDPPEPSPDPDPDSDSDSDSGSD----- 595
QY 742 KYTLKDENGNIISTTTTDENGKYQFDNLSGNVYVHFDPKPSMGTOTTTDSGDDEODADG 801
DB 536 -----SDSGSDSDSESD- 607
QY 802 BEVHYTIDHDFSIDNGYDESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 861
DB 608 -----SDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 660
QY 862 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 921
DB 661 DSDSDSDSDSESDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 720
QY 922 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 981
DB 721 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 780
QY 982 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 1041
DB 781 DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS 840
QY 1042 G-----NSDCKSTR-----DKLPDTGANEDYSGKTLGLTLAGLAL 1079
DB 841 RYTPPNEQKAPSNPKGEVHNHNSKVKQHKTDALPETG-DKSENTNATLFGAMMALLSGL 899
QY 1080 LL-GKRRKRRKKK 1091
DB 900 LLEFRKKQDHKEK 912
```

RESULT 15  
Q932C5

```
ID 0932C5 PRELIMINARY: PRT: 935 AA.
AC Q932C5.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibriogen-binding protein.
CN FNB OR S4V0811.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iken J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani U.I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Iabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RL Lancet 357:1225-1240(2001).
DR EMBL, AP003360; BAB56973.1; -.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMS: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00402; BPD_TRANSP_INN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;

Query Match 28.0%; Score 1581; DB 16; Length 935;
Best Local Similarity 36.7%; Pred. No. 4,4e-46;
Matches 412; Conservative 158; Mismatches 310; Indels 244; Gaps 36;

QY 18 KSNKYAIRKFTVGTASIVGATLLFG-LGNHAKAEENSVDQKSNTPDDELSDSDSS 76
DB 5 KKEKAIRKKSIGVASVLGTLIGGLSSKEADSENSV-----TQSDSASN 52
QY 77 DEKNDVINNNOSINTDNNQIITKEETNNNYGIEKRSDEKREESTTNDENATFLQKTP 136
DB 53 ESKSNDSSVSAAPKTDDTNV-----SDPKTSNTN--NGETSYAQNKP 93
QY 137 QNTHTEEEVSSSVSSSSSIDTAQOPSHHTINREESVOTSDNVEDSHVSDFANSKI 196
DB 94 -----AQOETTOSSSTNATTEETPYTGATITTTNQNATPATTOS-SNTNDELVN--- 143
QY 197 KESNTESGKEENTIEOPNKKVEDSTTSQPSGYTNIDEKISNODELLN-PINEYENKARP 255
DB 144 QTSNETTSNDTNTV-----SSVNSPONSNTNAENVSFTQDTEATPPSN--NESAP 191
QY 256 LSTTSA-----QPSIKRVTYNQJLA-----EGGSNNVHLIKVTOOSTIEGDDSE 300
DB 192 QNTDASNKDVSOAVNPSTPRKRAFSLAVALADAPAAAGTDITN--GLTVKYVT--IDS 246
QY 301 GYTKAHDAENLIYDVFEDDKVKSQDTMTVIDDKTVPDSDLTSTFIKIDNSGEIAT 360
DB 247 TIVYPHQAGYVKNLNFSPNSAVKADTFKITYPKLNLNGYTSIAKVPPIAAGD-QVILA 305
QY 361 TGTYDNKNKOITFTFDYDKYENIKAKHLKLTYSIDSKVPNNNTKL-DVEKTKAL--SS 417
DB 306 NGVIDS-DCGNVITFTFDYDNKENYNTANITMPAYID-----PENVTGTGNTVLTGIGTGN 360
QY 418 VNKTIIVEYORPNEFTANLQSMFTNIDTKNHTVEQTIYINPLRYSAKETNV-----N 470
DB 361 ASKTYLIDIEKYGOQFNHLSIKGTFIDQIDKTNNTYROTIIYVNP-----SGDNVVLPAITGN 415
QY 471 ISGNGDEGSTIIDSTIIVYKYGDNQNLPSNRIDYSEYEDVNTDVAOLGNNDVNI 530
DB 416 LIPNTKSNALIDAKNTDIAKRYV-DNANDLSESYIYNPSDFEDYTNQVATISPPNANQYKV 474
```





GenCore version 5.1.4.F5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:55 ; Search time 25.0323 Seconds

(without alignments)  
3098.077 Million cell updates/sec

Title: US-09-147-405b-13

Perfect score: 3013

Sequence: 1 SSDEKNDVYNNQSTINTD.....PGCSKSVRTDEGKIQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3013	100.0	1092	19	AAW41602
2	2810.5	93.3	930	23	ABP40469
3	2810.5	93.3	991	21	AAV83171
4	2810.5	93.0	991	21	AAV70120
5	1144.5	38.0	1166	20	AAV08643
6	857	28.4	278	22	AAV82803
7	576	19.1	1315	20	AAV08642
8	576	19.1	1349	22	AAU34402
9	576	19.1	1349	22	AAU37544
10	499.5	16.6	932	22	AAU34082

11	499.5	16.6	932	22	AAU36845	Staphylococcus aur
12	499	16.6	1802	21	AAV83170	Cell wall protein
13	499	16.6	1802	21	AAV70119	Staph. epidermidis
14	497	16.5	1155	22	AAV82343	S. epidermidis ope
15	455	15.1	841	22	AAU34283	Staphylococcus aur
16	455	15.1	841	22	AAU37158	Staphylococcus aur
17	455	15.1	930	20	AAV08641	S. aureus SdrC pro
18	454	15.1	936	18	AAV89801	Staphylococcus aur
19	452	15.0	933	21	AAV58435	Staphylococcus aur
20	452	15.0	933	22	AAV69508	Staphylococcus aur
21	449	14.9	1021	22	AAU33975	Staphylococcus aur
22	449	14.9	1021	22	AAU36951	S. aureus ClfB pro
23	396	13.1	918	20	AAV08640	S. aureus antigen
24	377	12.5	496	23	AAU75490	Staphylococcus aur
25	376	12.5	767	22	AAU34403	Staphylococcus aur
26	376	12.5	767	22	AAU37547	Staphylococcus aur
27	373	12.4	940	11	AAV07070	Fibronectin-bindin
28	357.5	11.9	238	18	AAV28019	Staphylococcus aur
29	349	11.6	345	19	AAV31555	Fibronectin-bindin
30	319.5	10.6	978	22	AAU33960	Staphylococcus aur
31	319.5	10.6	1001	22	AAU37093	Staphylococcus aur
32	308.5	10.2	1018	22	AAU34301	Staphylococcus aur
33	308.5	10.2	1018	22	AAU37245	Staphylococcus aur
34	300.5	10.0	1027	18	AAV89806	Staphylococcus aur
35	297.5	9.9	1018	9	AAV82115	Fibronectin bindin
36	293.5	9.7	970	23	ABP27418	Streptococcus poly
37	259.5	8.6	1112	20	AAV08603	S. pyogenes SFBP-
38	259.5	8.6	1161	23	AAE22273	Streptococcal fibr
39	220.5	7.3	3696	23	ABP40235	Staphylococcus epl
40	219	7.2	537	7	AAV60452	Sequence of the As
41	216.5	7.2	1959	23	ABV49411	Listeria monocytog
42	212.5	7.1	1279	22	AAV83047	S. epidermidis ope
43	211.5	7.0	2206	21	AAV18254	Plasmodium falcipa
44	209	6.9	570	23	AAU67000	Staphylococcus aur
45	208.5	6.9	251	18	AAV89804	Staphylococcus aur

#### ALIGNMENTS

RESULT 1

AAW41602 standard; Protein: 1092 AA.

AC AAW41602;

XX 22-JUN-1998 (first entry)

DE Staphylococcus epidermidis fibrinogen binding protein FIG.

XX Fibrinogen binding protein; FIG; aggregation; infection;

KW coagulase-negative Staphylococcus; therapy; diagnosis;

KW immunisation; immunogen; vaccine.

XX

OS Staphylococcus epidermidis strain HB.

XX

XX Key

XX Peptide

FT 1..51

FT /label- Sig-peptide

FT 52..1092

FT /label- Mat-protein

FT 52..824

FT /note- "non-repetitive region, harbours

FT fibrinogen binding activity"

FT Region

FT 825..1040

FT /note- "Asp-Ser dipeptide repeat region";

FT 1053..1057

FT /note- "cell wall anchoring motif"

XX

XX MO9748727-A1.

XX 24-DEC-1997.

PF 18-JUN-1997; 97MO-SE01091.  
 XX 20-JUN-1996; 96SE-0002496.  
 XX (FLOC/) FLOCK J.  
 PA (FRYK/) FRYKBERG L.  
 PA (GUSS/) GUSS B.  
 PA (LIND/) LINDBERG M.  
 PA (NILS/) NILSSON M.  
 XX Flock J, Frykberg L, Guss B, Lindberg M, Nilsson M;  
 XX MPI: 1998-063079/06.  
 DR N-PSDB: AAV04279.  
 XX  
 XX Fibrinogen-binding protein from coagulase-negative Staphylococcus  
 PT used for prevention, treatment and diagnosis of staphylococcus  
 PT infection  
 XX  
 XX Example 3; Fig 6; 45pp; English.  
 XX The protein comprises the fibrinogen binding protein (FIG) of  
 CC coagulase-negative Staphylococcus epidermidis HB. Its amino  
 CC acid sequence was deduced from the isolated fig gene (see AAV04279).  
 CC The closest known analogue of FIG is the clumping factor of  
 CC S. aureus which also binds fibrinogen and promotes bacterial  
 CC aggregation in serum. Recombinant FIG polypeptides can be  
 CC expressed in host cells. They are used as immunogens, particularly  
 CC in vaccines (which may be expressed in vivo) to protect humans and  
 CC animals against coagulase-negative Staphylococcus infection.  
 CC Antibodies raised against FIG can be used for passive immunisation.  
 CC They block the adherence of bacteria and for diagnosis.  
 XX  
 SQ Sequence 1092 AA;

Query Match 100.0%; Score 3013; DB 19; Length 1092;

Best Local Similarity 100.0%; Pred. No. 1.2e-156; Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDEEKNVYINNQSINTDDNNOI IKKETNNYDGIKRSDEPTESTTNVDEAFLOK 60  
 DB 75 SSDEKNVYINNQSINTDDNNOI IKKETNNYDGIKRSDEPTESTTNVDEAFLOK 134  
 QY 61 TPQDNHLEEEVKSSSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVSFPANS 120  
 DB 135 TPQDNHLEEEVKSSSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVSFPANS 194  
 QY 121 KIKESNTSGKEENTLEOPKRVKEDSTTSQPSGYTNIDEKISMODELNLPINEYENKAR 180  
 DB 195 KIKESNTSGKEENTLEOPKRVKEDSTTSQPSGYTNIDEKISMODELNLPINEYENKAR 254  
 QY 181 PLSTSAQPSIKRVYNOLAEGSNVNHILIKYTDOSITEGYDSEGVIAHAEMLIYD 240  
 DB 255 PLSTSAQPSIKRVYNOLAEGSNVNHILIKYTDOSITEGYDSEGVIAHAEMLIYD 314  
 QY 241 VFPEVDKYSKGTMTVDIDKNTVPSDLTDSFTPIKIDNSGFIITATGYDNKKOITYT 300  
 DB 315 VFPEVDKYSKGTMTVDIDKNTVPSDLTDSFTPIKIDNSGFIITATGYDNKKOITYT 374  
 QY 301 FTDYVYKYEIKAHKLITSYIDSKVPNNNTKLDVEYKTAALSSVNTITVEYORPENRT 360  
 DB 375 FTDYVYKYEIKAHKLITSYIDSKVPNNNTKLDVEYKTAALSSVNTITVEYORPENRT 434  
 QY 361 ANLOSMTNIDTNNHTVEQTIYINPLRYSAKETNVIISGNGDSTIIDSTIIKKYKYG 420  
 DB 435 ANLOSMTNIDTNNHTVEQTIYINPLRYSAKETNVIISGNGDSTIIDSTIIKKYKYG 494  
 QY 421 DNQNLDSNRITDYSEEDVTNDYIAQLGNNDVYNINFGNIDSPYIIKVIKSDPKKDY 480  
 DB 495 DNQNLDSNRITDYSEEDVTNDYIAQLGNNDVYNINFGNIDSPYIIKVIKSDPKKDY 554  
 QY 481 TTIGQTVTQTTINETGFRRTASYNNTIAFSTSSGOGGDDLPPEKTYKIGDYVWEDVVK 540

DB 555 TTIGQTVTQTTINETGFRRTASYNNTIAFSTSSGOGGDDLPPEKTYKIGDYVWEDVVK 614  
 QY 541 DGIQNTDNEKPLSNVLTLPDGTSKSVRTDEDKRYGDFG 582  
 DB 615 DGIQNTDNEKPLSNVLTLPDGTSKSVRTDEDKRYGDFG 656

# RESULT 2

ABP40469

ID ABP40469 standard; Protein; 930 AA.

AC ABP40469;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5314.

KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;

XX antibacterial; gene therapy.

OS Staphylococcus epidermidis.

PN US6380370-B1.

XX 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

PR 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

DR MPI: 2002-381255/41.

XX N-PSDB: ABN93014.

PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis

PS polypeptide, useful for diagnosing and treating bacterial infections -

XX Disclosure; SEQ ID 5314; 267bp; English.

CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading

CC frame (ORF) nucleic acid sequences which encode the amino acid sequences

CC given in ABP35124 to ABP37960. The S. epidermidis sequences have

CC antibacterial activity and can be used in gene therapy. The sequences

CC can also be used in the diagnosis and treatment of bacterial infections,

CC particularly S. epidermidis infections. The sequences can be used to

CC screen for compounds able to interfere with the S. epidermidis life

CC cycle or inhibit S. epidermidis infection.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

XX USPIO web site.

SQ Sequence 930 AA;

Query Match 93.3%; Score 2810.5; DB 23; Length 930;

Best Local Similarity 93.6%; Pred. No. 1.2e-145; Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SDEEKNVYINNQSINTDDNNOI IKKETNNYDGIKRSDEPTESTTNVDEAFLOK 60

DB 73 SSDEKNVYINNQSINTDDNNOI IKKETNNYDGIKRSDEPTESTTNVDEAFLOK 131

QY 61 TPQDNHLEEEVKSSSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVSFPANS 120

DB 132 TPQDNHLEEEVKSSSVSSNSIDTAQPSHTTINREESVQTSNDVEDSHVSFPANS 191

QY 121 KIKESNTSGKEENTLEOPKRVKEDSTTSQPSGYTNIDEKISMODELNLPINEYENKAR 180

DB 192 KIKESNTSGKEENTLEOPKRVKEDSTTSQPSGYTNIDEKISMODELNLPINEYENKAR 251

```

OY 181 PLSTTSAOPSSIKRVYNOLAEOGSNVNHLIKVTPDSTEGYDSEGYIKAHDAENLYD 240
DB 252 PLSTTSAOPSSIKRVYNOLAEOGSNVNHLIKVTPDSTEGYDSDGIIKAHDAENLYD 311
OY 241 VTFEVDKVKSGDPTMTVDIDKNTVPSDLTDSFTIPKIDNSSEIATGTYDNKNKOITYT 300
DB 312 VTFEVDKVKSGDPTMTVDIDKNTVPSDLTDSFAIPKIDNSSEIATGTYDNKNKOITYT 371
OY 301 FTDYVDKVENIKAHKLKLSYIDSKVPPNNNTKLDVEYKTAALSSVVKTTTVEYQRENERT 360
DB 372 FTDYVDKVENIKAHKLKLSYIDSKVPPNNNTKLDVEYKTAALSSVVKTTTVEYQRENERT 431
OY 361 ANLOSMFNTIDPKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVVYVG 420
DB 432 ANLOSMFNTIDPKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVVYVG 491
OY 421 DNQNLPSDSNRIDYSEYEDVTNDYVLAQLGNNNDVNIENFGNIDSPYIIKIVISKYDPNKDDY 480
DB 492 DNQNLPSDSNRIDYSEYEDVTNDYVLAQLGNNNDVNIENFGNIDSPYIIKIVISKYDPNKDDY 551
OY 481 TTIOQVTVMQTTINEYTGEPRTASYDNTIAFSTSSGOGGDLPEPKTYIKIDYVWEDVDK 540
DB 552 TTIOQVTVMQTTINEYTGEPRTASYDNTIAFSTSSGOGGDLPEPKTYIKIDYVWEDVDK 611
OY 541 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTDEGKYOFDG 582
DB 612 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTDEGKYOFDG 653

RESULT 3
AA83171
ID AA83171 standard; Protein: 991 AA.
AC AA83171;
XX
DT 24-JUL-2000 (first entry)
XX
DE Cell wall protein Sdrg.
XX
KV Sdrg; Sdrg; Sdrg; coagulase negative; staphylococcus; scepticemia;
KW osteomyelitis; endocarditis; immune response; vaccine; graft;
KW stent; intravenous catheter; heart valve; cardiac.
XX
OS Staphylococcus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 14
FT /note= "Position encoded by TAG stop codon"
FT Misc-difference 33
FT /note= "Position encoded by TGA stop codon"
FT Misc-difference 964
FT /note= "Position encoded by TAA stop codon"
FT /note= "Position encoded by TAA stop codon"
FT Misc-difference 980
FT /note= "Position encoded by TAG stop codon"
FT /note= "Position encoded by TAG stop codon"
FT Misc-difference 989
FT /note= "Position encoded by TAA stop codon"
XX
PN WO200012689-A1.
XX
PD 09-MAR-2000.
XX
PF 31-AUG-1999; 99WO-US19728.
XX
PR 31-AUG-1998; 98US-0098443.
PR 25-JAN-1998; 99US-0117119.
XX
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
PA (TEXA ) UNIT TEXAS A & M SYSTEM.
XX
XX Foster TJ, Hook M, Davis S, Hartford O, McCrea K, Ni Eldhin D;
XX
XX MPI: 2000-256637/22.
XX
XX N-PSDB; AA293534.

```

```

XX
XX Recombinant or synthetic proteins from coagulase-negative staphylococci
PT useful for prevention, treatment and diagnosis of staphylococcal
PT infections bind soluble and immobilized fibrinogen
PS
PS Claim 8; Figure 3; 104pp; English.
XX
XX Isolated staphylococcus sdr cell wall proteins which bind both
CC soluble and immobilized fibrinogen are useful for treating or
CC preventing coagulase-negative staphylococcal infection such as
CC scepticemia, osteomyelitis or endocarditis, and for inducing immune
CC responses in patients. The cell wall proteins are also useful for
CC reducing coagulase-negative staphylococci infection of indwelling
CC medical devices such as vascular grafts, vascular stents,
CC intravenous catheters, artificial heart valves and cardiac assist
CC devices. The cell wall associated proteins are able to inhibit
CC staphylococcal adhesion to immobilised extracellular matrix or host
CC cells present on the surface of implanted biomaterials.
XX
XX Sequence 991 AA:
XX
XX Query Match 93.3%; Score 2810.5; DB 21; Length 991;
XX Best Local Similarity 93.6%; Pred. No. 1.3e-145;
XX Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;
OY 1 SDEEKNVYINNNOISINDNNQIIKKEETNNYDIEKRSDERTSTINVDENATFLOK 60
DB 106 SNEEKNVYINNNOISINDNNQ-IKKEETNSDAIENSKDITOSTINVDENATFLOK 164
OY 61 TPQDNTHLTEEEKSSSESSSSSIDTAAQOQSHTTINREESVOTSDNVEDSHVDFANS 120
DB 165 TPQDNTOLKEEYVKEPSSSVSSSSMDTAAQOQSHTTINSEASIQSDNEENRVDFANS 224
OY 121 KIKESNTESGKEENTIEOPNKVKESTTSQPSGYTNIDEKISNOBELNLPINEYENKAR 180
DB 225 KIEESNTESKKEENTIEOPNKVREDSITSQPSYKNIDEKISNOBELNLPINEYENKAR 284
OY 181 PLSTTSAOPSSIKRVYNOLAEOGSNVNHLIKVTPDSTEGYDSEGYIKAHDAENLYD 240
DB 285 PLSTTSAOPSSIKRVYNOLAEOGSNVNHLIKVTPDSTEGYDSDGIIKAHDAENLYD 344
OY 241 VTFEVDKVKSGDPTMTVDIDKNTVPSDLTDSFTIPKIDNSSEIATGTYDNKNKOITYT 300
DB 345 VTFEVDKVKSGDPTMTVDIDKNTVPSDLTDSFAIPKIDNSSEIATGTYDNKNKOITYT 404
OY 301 FTDYVDKVENIKAHKLKLSYIDSKVPPNNNTKLDVEYKTAALSSVVKTTTVEYQRENERT 360
DB 405 FTDYVDKVENIKAHKLKLSYIDSKVPPNNNTKLDVEYKTAALSSVVKTTTVEYQRENERT 464
OY 361 ANLOSMFNTIDPKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVVYVG 420
DB 465 ANLOSMFNTIDPKNHTVEQTIYINPLRYSAKETNNVINSNGDEGSTIIDSTIIKVVYVG 524
OY 421 DNQNLPSDSNRIDYSEYEDVTNDYVLAQLGNNNDVNIENFGNIDSPYIIKIVISKYDPNKDDY 480
DB 525 DNQNLPSDSNRIDYSEYEDVTNDYVLAQLGNNNDVNIENFGNIDSPYIIKIVISKYDPNKDDY 584
OY 481 TTIOQVTVMQTTINEYTGEPRTASYDNTIAFSTSSGOGGDLPEPKTYIKIDYVWEDVDK 540
DB 585 TTIOQVTVMQTTINEYTGEPRTASYDNTIAFSTSSGOGGDLPEPKTYIKIDYVWEDVDK 644
OY 541 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTDEGKYOFDG 582
DB 645 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTDEGKYOFDG 686

RESULT 4
AA70120
ID AA70120 standard; Protein: 991 AA.
AC AA70120;
XX
XX 06-JUN-2000 (first entry)

```

XX Staph. epidermidis serine-aspartate repeat region protein Sdrg.  
 DE  
 XX  
 KW Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAM;  
 KW microbial surface components recognising adhesive matrix molecules;  
 KW collagen binding protein; CBP; CNA; fibrinogen binding protein;  
 KW Clumping factor A; ClfA; Clumping factor B; ClfB; FBP;  
 KW fibrinectin binding protein; Staphylococcus infection;  
 KW serine-aspartate repeat region protein; SDR protein; Sdrg.  
 XX  
 OS Staphylococcus epidermidis.  
 FH  
 FH Key Location/Qualifiers  
 FT MISC-difference 14  
 FT MISC-difference 33 /note- "Encoded by in-frame stop codon TAG"  
 FT MISC-difference 964 /note- "Encoded by in-frame stop codon TGA"  
 FT MISC-difference 980 /note- "Encoded by in-frame stop codon TAA"  
 FT MISC-difference 989 /note- "Encoded by in-frame stop codon TAG"  
 FT MISC-difference 989 /note- "Encoded by in-frame stop codon TAA"  
 PN W0200012131-A1.  
 PN 09-MAR-2000.  
 PF 31-AUG-1999; 99MO-US19727.  
 PR 31-AUG-1998; 98US-0098439.  
 PA (INH1-) INHIBITEX INC.  
 PA (TEXA-) UNIV TEXAS A & M SYSTEM.  
 PA (QUE-) QUEEN ELIZABETH COLLEGE DUBLIN.  
 PI Patti JM, Foster TJ, Hook M;  
 XX WPI: 2000-237781/20.  
 DR N-PSDB; AA251202.  
 XX  
 PT Composition used for generating immune response or for inhibiting  
 PT microbial colonization in an animal comprises antibodies that bind  
 PT collagen binding protein, fibrinogen binding protein and, optionally,  
 PT fibronectin binding protein  
 XX  
 YY Claim 8; Fig 4; 115pp; English.

CC The patent discloses multicomponent vaccines containing selected  
 CC combinations of bacterial binding proteins termed MSCRAM (microbial  
 CC surface components recognising adhesive matrix molecules) or their  
 CC antibodies. A vaccine composition is provided that includes collagen  
 CC binding protein or peptide, e.g. CNA, a fibrinogen binding protein  
 CC preferably Clumping factor A (ClfA) or Clumping factor B (ClfB),  
 CC and optionally a fibronectin binding protein e.g. FBP-A.  
 CC The vaccines are useful for imparting protection against a broad  
 CC spectrum of Staphylococcal strains and for inhibiting microbial  
 CC colonisation, especially of Staphylococcus aureus, in an animal.  
 CC The combinations can also be used to select donor blood pools for the  
 CC preparation of purified blood products for passive immunisation.  
 CC The present sequence is a serine-aspartate repeat region  
 CC protein, Sdrg from Staphylococcus epidermidis. The Sdr protein is  
 CC useful in vaccine preparation in combination with specific  
 CC bacterial binding proteins. These vaccines can be used to treat a broad  
 CC spectrum of bacterial infections, including those arising from both  
 CC coagulase-positive and coagulase-negative bacteria.  
 CC  
 XX Sequence 991 AA;

Query Match 93.38; Score 2810.5; DB 21; Length 991;  
 Best Local Similarity 93.68; Pred. No. 1.3e-145;  
 Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEKNDVNNOSINTDDNNQIIKKETNYDIEKRSRDESTTNDENAEFLQK 60  
 DB 106 SSNEKNDVNNOSINTDDNO-IRKEETNSDAIEKRSRSDITOSTNDENAEFLQK 164  
 QY 61 TPQDNTLITEEVSESSVSSNSIPTAQPSTTTINRESVOTSDNVEDSHSDPANS 120  
 DB 165 TPQDNTLKEEVAVPESSVESNSMPTAQPSHTTINSEASITSDNEERSRSDPANS 224  
 QY 121 KIKESNTESGKEENTIQPNKVKEDSTTSQPSGYTINDEKISNDELINLPINYEKAR 180  
 DB 225 KIIESNTESKKEENTIEQPNKVEDSTTSQPSGYTINDEKISNDELINLPINYEKAR 284  
 QY 181 PLSTTSAPSKRYTVNQLAEOGSNNHLIKYVDOSTTEGYDSDGTYRAHDAENLYD 240  
 DB 285 PLSTTSAPSKRYTVNQLAEOGSNNHLIKYVDOSTTEGYDSDGTYRAHDAENLYD 344  
 QY 241 VTFEVDKVKSGDTMTYDIDKNTVPSDLTDSFTIPKIKDNGSELIATGYDNKKQITYT 300  
 DB 345 VTFEVDKVKSGDTMTYDIDKNTVPSDLTDSFALPKIKDNGSELIATGYDNKKQITYT 404  
 QY 301 FTDYVDKYEENIKAKLITSTIDSKVPNNNTKLDVEYKKTALSSVNTTVEYQRPENRT 360  
 DB 405 FTDYVDKYEENIKAKLITSTIDSKVPNNNTKLDVEYKKTALSSVNTTVEYQRPENRT 464  
 QY 361 ANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDESGTIIIDSTIIKVKYVG 420  
 DB 465 ANLOSMTNIDTKNHTVEQTIYINPLRYSAKETNVIISNGDESGTIIIDSTIIKVKYVG 524  
 QY 421 DNQNLPSNRIYDYSEVEDYTNDDYAOLGNNNDVNIENFGNIDSPYIIKIVISKYDPNNDY 480  
 DB 525 DNQNLPSNRIYDYSEVEDYTNDDYAOLGNNNDVNIENFGNIDSPYIIKIVISKYDPNNDY 584  
 QY 481 TTIOQTYTMOITINIEYGEERTASYDNTIAFSTSSGOGGDLPEPKYTKIGDYWEVDK 540  
 DB 585 TTIOQTYTMOITINIEYGEERTASYDNTIAFSTSSGOGGDLPEPKYTKIGDYWEVDK 644  
 QY 541 DGIONTNDNEKPLSNVLTLYTPGTSKSVRTDDEGKYOPDG 582  
 DB 645 DGIONTNDNEKPLSNVLTLYTPGTSKSVRTDDEGKYOPDG 686

RESULT 5  
 AAY08643  
 ID AAY08643 standard; protein; 1166 AA.  
 XX  
 AC AAY08643;  
 XX  
 DT 09-AUG-1999 (first entry)  
 XX  
 DE S. aureus Sdrg protein.  
 XX  
 KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; Sdrg; Sdrg;  
 KW Sdrg; fibrinogen; medical device; competitive inhibitor; pharmaceutical;  
 KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;  
 KW extracellular matrix; vascular graft; vascular stent; vaccine;  
 KW intravenous catheter; artificial heart valve; cardiac assist device;  
 KW antibacterial.  
 KW  
 OS Staphylococcus aureus.  
 XX  
 PN MO9927109-A2.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PF 25-NOV-1998; 98MO-US25246.  
 XX  
 PR 31-AUG-1998; 98US-0098427.  
 XX  
 PR 26-NOV-1997; 97US-0066815.  
 XX  
 PA (EIDH/) EIDHIN D N.  
 PA (FORF-) FORAS T/A BIORESEARCH IRELAND.  
 PA (FOST/) FOSTER T J.  
 PA (HOOK/) HOOK M A O.

PA (INHIT-) INHIBITEX INC.  
 PA (JOSEF) JOSEFSSON E.  
 PA (PATTI) PATTI J M.  
 PA (PERK) PERKINS S E.  
 XX  
 PI Eldin DN, Foster TJ, Hook MAO, Josefsson E, Patti JM:  
 PI Perkins SE;  
 XX WPI: 1999-357844/30.  
 DR N-PSDB: AAX77594.  
 XX  
 PT Staphylococcus aureus fibrinogen-binding proteins for treating  
 PT septicemia, osteomyelitis, mastitis or endocarditis  
 XX  
 PS Claim 8; Fig 9; 143pp; English.  
 XX  
 CC This invention describes novel Staphylococcus aureus fibrinogen-binding  
 CC proteins that bind both the alpha and beta fibrinogen chains. The  
 CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and  
 CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere  
 CC to medical devices, binding proteins that bind both the alpha and beta  
 CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as  
 CC competitive inhibitors to block this binding. Antibodies against ClfB,  
 CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.  
 CC The proteins of the invention can be used in a pharmaceutical composition  
 CC for the treatment of Staphylococcus aureus infection e.g. septicemia,  
 CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of  
 CC S. aureus to the extracellular matrix. The proteins or their fragments  
 CC may be used to coat a medical device to reduce the S. aureus infection of  
 CC an indwelling medical device, especially where the medical device is  
 CC selected from the group consisting of vascular grafts, vascular stents,  
 CC intravenous catheters, artificial heart valves, and cardiac assist  
 CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or  
 CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat  
 CC region or a gene encoding it may be used as an identifying probe for the  
 CC identification of genes and encoding proteins from Staphylococcus aureus  
 CC (other than ClfA), S. hemolyticus, S. lugdunensis, and S. schleiferi.  
 CC The proteins of the invention have antibacterial activity.  
 CC  
 XX  
 SQ Sequence 1166 AA;

Query Match 38.0%; Score 1144.5; DB 20; Length 1166;  
 Best Local Similarity 42.6%; Pred. No. 2e-54;  
 Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;

QY 1 SDEKNDYINNNOISINTDNNQI--IKKEFNINNYGDIKRSRDRRESTTNDENATFL 58  
 DB 66 ATTSDKREVVSSTENNSTNNSTNPKE--TNTDSQPEAKKESTSSSTOKQNNVTAF 123  
 QY 59 OKTPDNTLHTEBEVES--SSVESNSSIDTAQOPSHTTTIRRESVOTSDVNESSHVS 115  
 DB 124 TETKPN--IEKENVAPSTDKATETDTSVILEKKAPNNT--NNDVTPRS-----TS 172  
 QY 116 DEANSKIKESNT--ESGKEENTIEQPNKVKEDSTTSQPSGYTINDEKIS-NODELLNLP 171  
 DB 173 ESTSTBIQKPTTPOESTNINENSOQPPPTSKVD--NOVTDATNPREPVNSKEELKNP 229  
 QY 172 -----INEYENKARPLSTSAOPSIKRV-----TYNQLAAAGGSVNNLIVTQD 216  
 DB 230 EKLKELVRNDSNTHSTKPAATPATSAPRVNAKMRFAVAOPAAVAASNNVNLIVTKG 289  
 QY 217 SITEGYDSEGVKKAHDAENLIVDFEVDDKSGDTMVDIDKTKPSPDLDSFTIPK 276  
 DB 290 TIVAG-DGKDNVAHAHDGKDIEDTEFTIDNKYKKGDTMTINDKNIVPSLDLKDNDPID 348  
 QY 277 IKDNGSEIIATGTGRKKNQIYTFDYVDKYENIKAHKLKTSYIDSKVPNNNTKLVE 336  
 DB 349 ITDPSEVIKAKTGFDAKQKITTYTFDYVDKYEDIKSRLLTLYSIDKKVTP-NETSLNT 407  
 QY 337 YKALSSVKTITVEYQRPENNTANLQSMFTNIDRKNHVEQDTYINPLRYAKETNNV 396  
 DB 408 FATAGKETSONVTVDYODPVAVHDSNOSTLFTKLEDEKQIEQIYVNNPLKSAATNTKVD 467

QY 397 ISGNG-----DEGSTIIDSTIIKRYKYGDNQNPDSNRITVSESEEDVTND-DYA 446  
 DB 468 IAGSQVDYDGNIGKILGNSITIDONTIEIKYKNSDQQLPQSRRTVDFSOYEDVSOFPNK 527  
 QY 447 OLGNNNDVINENGIDSPYIIKIVISKYDPNKKDYTTIQOTVMTQITNEYTGEEFTASVD 506  
 DB 528 KFSNNVATLDPGDIINSAYIIKIVSKYTPTSIGELDIAQGSMTRT-DKY-GYIVYAGYS 585  
 QY 507 NTIAFSTSSGQGDLL-PPEKTYKIGDYVEDVDKGIQNTNDNKPSPLSNVLVLTPTPDG 565  
 DB 586 NFIVTSNDPTGGDGTVPKEEKLYKIGDYVEDVDKDGVGQFTSKERKPMANVLVLTPTPDG 645  
 QY 566 TSKSVRTDEDKTYQPDG 582  
 DB 646 TTKSVRTDANGHYEFGG 662

RESULT 6  
 AAG82803  
 ID AAG82803 standard; Protein: 278 AA.  
 XX  
 AC AAG82803;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX  
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:2700.  
 XX  
 OS Staphylococcus epidermidis SRI strain; infection; diagnosis;  
 XX  
 KW vaccination; endocarditis.  
 XX  
 PN Staphylococcus epidermidis.  
 XX  
 PD WO200134809-A2.  
 XX  
 PF 17-MAY-2001.  
 XX  
 PR 09-NOV-2000; 2000WO-US30782.  
 XX  
 PR 09-NOV-1999; 99US-0164258.  
 XX  
 PA (GLAX) GLAXO GROUP LTD.  
 XX  
 PI Kilmerry WI;  
 XX  
 DR WPI: 2001-316495/33.  
 XX  
 PT N-PSDB: AAH53653.  
 XX  
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
 XX  
 PS useful for vaccinating against infections, e.g. endocarditis -  
 XX  
 PS Claim 18; Page 705; 2188pp; English.  
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAH81454 to AAG83120, from Staphylococcus epidermidis.  
 CC (I) and (II) can have antibacterial activity and therefore can be used  
 CC in vaccination. The nucleic acids (I) may be used to produce the  
 CC S. epidermidis polypeptides (II) via the production of vectors  
 CC containing them which are used to produce hosts cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 CC used to vaccinate subjects and to raise antibodies against the bacteria.  
 CC The polypeptides may also be used to assay for other inhibitors of their  
 CC activity and therefore identify compounds that may be used for the  
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA  
 CC polynucleotide sequences from the present invention. AAH55091 to  
 CC AAH55098 represent oligonucleotide sequences and primers which are used  
 CC in the exemplification of the present invention.  
 CC N.B. The present invention specifically claims all the polynucleotide  
 CC sequences given in the sequence listing of the present specification,  
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,  
 CC no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 278 AA;  
Query Match 28.4%; Score 857; DB 22; Length 278;  
Best Local Similarity 93.0%; Pred. No. 1.6e-39;  
Matches 172; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 SDEKNDVNNNSINTDNNQIKKEETNNYDIEKRSDEPSTNVDENATFLOK 60  
DB 75 SSDEENOVNNNSINDNNQIKKEETNNNSDIEKRSDEPSTNVDENATFLOK 134  
QY 61 TPQDNTHTLEEVKSSSVSSSSSIDTAAQPSHTTINRESVQTSNVEDSHVSDFANS 120  
DB 135 SPQDNTHTLEEVKSSSVSSSSSIDTAAQPSHTTINRESVQTSNVEDSHVSDFANS 194  
QY 121 KIKESNTSGKEENTIEOPNKVKEDSTTSOPSGYTNIDEKISNDELNLPIENENKAR 180  
DB 195 KIKESNTSGKEENTIEOPNKVKEDSTTSOPSGYTNIDEKISNDELNLPIENENKAR 252  
181 PLSTP 185  
253 KLSVT 257

RESULT 7  
AY08642  
ID AY08642 standard; Protein: 1315 AA.  
XX  
AC AY08642;  
XX  
DT 09-AUG-1999 (first entry)  
XX  
DE S. aureus Sdrd protein.  
XX  
KW Fibrinogen-binding protein; alpha chain; beta chain; ClfB; SdrC; SdrD;  
KW SdrE; fibrinogen; medical device; competitive inhibitor; pharmaceutical;  
KW treatment; infection; septicemia; osteomyelitis; mastitis; endocarditis;  
KW extracellular matrix; vascular graft; vascular stent; vaccine;  
KW intravenous catheter; artificial heart valve; cardiac assist device;  
KW antibacterial.  
XX  
OS Staphylococcus aureus.  
XX  
PN MO9927109-A2.  
XX  
PD 03-JUN-1999.  
XX  
VX 25-NOV-1998; 98MO-US25246.  
PR 31-AUG-1998; 98US-0098427.  
XX 26-NOV-1997; 97US-0066815.  
XX  
PA (EIDH/) EIDHIN D N.  
PA (FORF-) FORFAS T/A BIORESEARCH IRELAND.  
PA (FOST/) FOSTER T J.  
PA (HOOK/) HOOK M A O.  
PA (INH-) INHIBITEX INC.  
PA (JOSE/) JOSESSON E.  
PA (PAT/) PATTI J M.  
PA (PERK/) PERKINS S E.  
XX  
PI Elidh D N, Foster T J, Hook M A O, Josefsen E, Patti J M;  
PI Perkins S E;  
XX  
DR WPI: 1999-357844/30.  
XX N-PSDB: AAX77593.  
XX  
PT Staphylococcus aureus fibrinogen-binding proteins for treating  
XX septicemia, osteomyelitis, mastitis or endocarditis  
XX  
XX Claim 8; Fig 8; 143pp; English.  
XX  
CC This invention describes novel Staphylococcus aureus fibrinogen-binding  
CC proteins that bind both the alpha and beta fibrinogen chains. The

CC proteins (and their encoding nucleic acids are ClfB, SdrC, SdrD and  
CC SdrE). Staphylococcus aureus is thought to utilize fibrinogen to adhere  
CC to medical devices, binding proteins that bind both the alpha and beta  
CC fibrinogen chains (ClfB, SdrC, SdrD and SdrE) can therefore be used as  
CC competitive inhibitors to block this binding. Antibodies against ClfB,  
CC SdrC, SdrD and SdrE inhibit ClfB, SdrC, SdrD and SdrE mediated binding.  
CC The proteins of the invention can be used in a pharmaceutical composition  
CC for the treatment of Staphylococcus aureus infection e.g. septicemia,  
CC osteomyelitis, mastitis or endocarditis or to inhibit the binding of  
CC S. aureus to the extracellular matrix. The proteins or their fragments  
CC may be used to coat a medical device to reduce the S. aureus infection of  
CC an indwelling medical device, especially where the medical device is  
CC selected from the group consisting of vascular grafts, vascular stents,  
CC intravenous catheters, artificial heart valves, and cardiac assist  
CC devices. ClfB, SdrC, SdrD, SdrE, or an active fragment, subdomain or  
CC encoding gene may be used as a vaccine. The DS (aspartate serine) repeat  
CC region or a gene encoding it may be used as an identifying probe for the  
CC identification of genes and encoding proteins from Staphylococcus aureus  
CC (other than ClfA), S. hemolyticus, S. lugdenensis, and S. schleiferi.  
CC The proteins of the invention have antibacterial activity.

Sequence 1315 AA;  
Query Match 19.1%; Score 576; DB 20; Length 1315;  
Best Local Similarity 30.1%; Pred. No. 2.9e-23;  
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 1 SDEKNDVNNNSINTDNNQIKKEETNNYDIEKRSDEPSTNVDENATFLOK 56  
DB 55 STNKLNEATTSASDNDQSDVDKQMLQNEEDNTKNDQKEMVSSQGNETTNGKMLIKE 114  
QY 57 FLOKTPQDNTHTLEEVKSSSVSSSSSIDTAAQPSHTTINRESVQTSNVEDSHVSDFANS 116  
DB 115 SVQSTGKVKVSTAKSDEQASPKSTNEDLNTKQ---TISNDEALQ-PDLQEKSVYN 168  
QY 117 FANSKIKESNTSGKEENTIEOPNKVKEDSTTSOPSGYTNIDEKISNDELNLPIENENKAR 173  
DB 169 ---VOPTNEENKKVD-----AKTEST-----LWKSADIKNDELTVLDNNSNS 209  
QY 174 EYENKARPLSTSAOP---SIRKVTYNQALAEQGSVNHILIKYTDGISTEGYDSECVI 229  
DB 210 NNEENNADILILPKSTAPKRLNTRMRAIAVQSPSTKAKNNDLITSNTPLTVADADKNNKIV 269  
QY 230 KAHDAENILYVTFEVDKVKSGDTMTVDIDKNY-----PSDLTDSFTLPKIKD-NSG 282  
DB 270 PAQDYLSLKSQIT--VDDKVKSGDYFTIKY--SDTYQVYGLNPEDEKN--IGDIKDPNNG 323  
QY 283 ELIATGTNDKNNKQITFTYDVKYENIKAKLITSYIDSKVYPNNNTKLDVEYKATLS 342  
DB 324 ETIATAKHDTANMLITFTYDVRFNVSVMGAINSTIMDADTIP--VSKNDVEFNVTIG 381  
QY 343 SVNKTITVEYQRPN--ENRTANLQSMFTNIDTKNHT-----VEQTYIYNPLRYSAKE 392  
DB 382 NTTTKRTANIQYPPDYVNEKNSIGSAFT--ETVSHVGNKKNPGYKQIYVNPSENSTLN 439  
QY 393 TNNVNI-----SGNGDEGSTIIDDSITIKYKYGDNQNLPSNRIDYS--EDEDVTNDYA 446  
DB 440 AKLVQAVHSSYPNNIGQINKVDIKIYQVPGYTL--NKGYDVNTRKELTDVYN-OYL 495  
QY 447 Q---LGNNDVNNNSINTDNNQIKKEETNNYDIEKRSDEPSTNVDENATFLOK 503  
DB 496 OKITTYGDNNSAVIDFGNADSAVYVWNTKRFQYTSSESTLVQMATLSTGN-----KSV 549  
QY 504 SYDNTIASTSSGQGGDLPEPKTYKIDYVWEDVDKDGIGONTDNEKPLSNVLTLYTP 563  
DB 550 STGNALGFTNNQSGAG---QEVYKIGINYYWEDNKNKGQVEL--GEKGVGNVTVTV-FD 602  
QY 564 DGTSKSVR---TDEGKY 578  
DB 603 NNTNTKVGCAVTKEDGYS 620

RESULT 8







```

OY 35 GIEKRSDEPRTSTNVDNEATFLQKTPQDNTHLTHEEVKSSSVSSNSIDPAQPSH 94
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 45 GHEAKAAEHNTGELNOSKNETT-----APSEN-----KTEKVDNRQO--- 82
OY 95 TTINRESVQTSNDVEDSHVDFANSKIKESNTSEKGEENTTEOPNKVKEDSTISQSGY 154
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 83 -----NNIEOSTSD--QPKVNESDNTSVKE--TTEEP---QNTTSQTPKQ 122
OY 155 TNIDKISNDELNLPINEYENKARPLSTSAOPSIK-----RYVNOQLAA-EGGSNVN 208
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 123 NN--DAMANKD---NLAQNISTQAKDVSTTPKTTTIKPRILNRAVNTVAAPQGGTVN 177
OY 209 HLKIVTDOST-----TEGYDSEGVIRKAHAENLIYDVEVDKVKSGDTM 255
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 178 DKVHFTNIDIALDKGHVKNKTGTGTEFWATSSDVLK-----LKANTIDDSVEKGTJF 229
OY 256 TVDIDKNTVPSDLTDSFTIKIKNSGEIATGTVDKNNKOITFTPDYVCKYENIKAKHL 315
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 230 TFKYQYFRPGSVRLPSQOTONLYNAQGNITIAKGIYDSESTTFTYTFVYDQYTNISGSF 289
OY 316 KLTGYIDSKVPPNNNTKLDVEYKTAALSSVNTKTIYEQRPENRPTANLOSMTNIDFKNH 375
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 290 EQVAFARENATTDKTAAPMEVTLGNDKYSKNVIYD---GNQKQQLISSTNYINNEDL 346
OY 376 TVEQTIYIN-PLRYSAKETNV-NISGNGDEGTTIIDSTIIKVVYKVGDNQ-----LPD 427
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 347 SRNMTVVYVNOQPKTYTKETFTVNTLT-----GKFNPDPAKNFKIYEV-TNQNQFVDSFTPD 400
OY 428 SNRIYDSEYEDVTDNDYAOAGNNNDVNIING--NIDSPYIIKIVSKYDPKDKDYTTIQ 485
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 401 TSKLTDVTDKFKIT---YSNDKTAIVDLNGOSSSDKOYIIQVAYPDSNSTDNGKIDY 457
OY 486 TVTMOQTINNEYGEFFRTASDNTIAFSTSSGOGGDLPEPKTYKIGDYVWEDVDKGIQ 545
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 458 TLETONKSSMSN-----SYSNVNGSSSTANGD-----QKKYNIAGDYVWEDVTNKGQD 505
OY 546 TNDNEKPLSNLVLTLYTPDGTG-KSVRTDEDGKYQFDG 562
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 506 A-NEKGIGKYVILKDSNGKELDRTTDENGKYOFTG 541

```

```

PI Haselbeck R, Ohlsen KL, Zyskind JM, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS54704.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12438; 511bp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the
XX genes, their use in the discovery of novel antibiotics, the essential
XX genes themselves and the encoded proteins. The prokaryotes used are
XX Escherichia coli, staphylococcus aureus, Salmonella typhi, Klebsiella
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX invention is also useful for the identification of potential new targets
XX for antibiotic development. The antisense nucleic acids can also be used
XX to identify proteins used in proliferation, to express these proteins,
XX and to obtain antibodies capable of binding to the expressed proteins.
XX The proteins can be used to screen compounds in rational drug discovery
XX programmes. The antisense nucleic acid sequence is also useful to screen
XX for homologous nucleic acids which are required for cell proliferation in
XX a wide variety of organisms. The present sequence represents an
XX essential prokaryotic cellular proliferation protein.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 932 AA;
XX
XX Query Match 16.6%; Score 499.5; DB 22; Length 932;
XX Best Local Similarity 27.7%; Pred. No. 2.8e-19;
XX Matches 160; Conservative 95; Mismatches 212; Indels 11; Gaps 24;
OY 35 GIEKRSDEPRTSTNVDNEATFLQKTPQDNTHLTHEEVKSSSVSSNSIDPAQPSH 94
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 45 GHEAKAAEHNTGELNOSKNETT-----APSEN-----KTEKVDNRQO--- 82
OY 95 TTINRESVQTSNDVEDSHVDFANSKIKESNTSEKGEENTTEOPNKVKEDSTISQSGY 154
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 83 -----NNIEOSTSD--QPKVNESDNTSVKE--TTEEP---QNTTSQTPKQ 122
OY 155 TNIDKISNDELNLPINEYENKARPLSTSAOPSIK-----RYVNOQLAA-EGGSNVN 208
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 123 NN--DAMANKD---NLAQNISTQAKDVSTTPKTTTIKPRILNRAVNTVAAPQGGTVN 177
OY 209 HLKIVTDOST-----TEGYDSEGVIRKAHAENLIYDVEVDKVKSGDTM 255
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 178 DKVHFTNIDIALDKGHVKNKTGTGTEFWATSSDVLK-----LKANTIDDSVEKGTJF 229
OY 256 TVDIDKNTVPSDLTDSFTIKIKNSGEIATGTVDKNNKOITFTPDYVCKYENIKAKHL 315
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 230 TFKYQYFRPGSVRLPSQOTONLYNAQGNITIAKGIYDSESTTFTYTFVYDQYTNISGSF 289
OY 316 KLTGYIDSKVPPNNNTKLDVEYKTAALSSVNTKTIYEQRPENRPTANLOSMTNIDFKNH 375
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 290 EQVAFARENATTDKTAAPMEVTLGNDKYSKNVIYD---GNQKQQLISSTNYINNEDL 346
OY 376 TVEQTIYIN-PLRYSAKETNV-NISGNGDEGTTIIDSTIIKVVYKVGDNQ-----LPD 427
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 347 SRNMTVVYVNOQPKTYTKETFTVNTLT-----GKFNPDPAKNFKIYEV-TNQNQFVDSFTPD 400
OY 428 SNRIYDSEYEDVTDNDYAOAGNNNDVNIING--NIDSPYIIKIVSKYDPKDKDYTTIQ 485
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 401 TSKLTDVTDKFKIT---YSNDKTAIVDLNGOSSSDKOYIIQVAYPDSNSTDNGKIDY 457
OY 486 TVTMOQTINNEYGEFFRTASDNTIAFSTSSGOGGDLPEPKTYKIGDYVWEDVDKGIQ 545
    | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 458 TLETONKSSMSN-----SYSNVNGSSSTANGD-----QKKYNIAGDYVWEDVTNKGQD 505

```

QY 546 TNDNEKPLSNVLTLYPDGTS-KSVTDEDEKGYQFDG 582  
 DB 506 A--NEKGKGVYVLKDSNGEKELDRITTDENGKGYQFTG 541

# RESULT 12

AY83170  
 ID AY83170 standard; Protein; 1802 AA.

AY83170;

24-JUL-2000 (first entry)

Cell wall protein SdrF.

SdrF; SdrF; coagulase negative; staphylococcus; septicemia;  
 osteomyelitis; endocarditis; immune response; vaccine; graft;  
 stent; intravenous catheter; heart valve; cardiac.

Staphylococcus sp.

Key Location/Qualifiers

FT MISC-difference 12 /note- "Position encoded by TAA stop codon"

FT MISC-difference 28 /note- "Position encoded by TAA stop codon"

FT MISC-difference 1771 /note- "Position encoded by TAG stop codon"

FT MISC-difference 1774 /note- "Position encoded by TAG stop codon"

FT MISC-difference 1793 /note- "Position encoded by TGA stop codon"

FT MISC-difference 1800 /note- "Position encoded by TAA stop codon"

FT MISC-difference 1800 /note- "Position encoded by TGA stop codon"

PN MO200012689-A1.

PD 09-MAR-2000.

PF 31-AUG-1999; 99MO-0519728.

PR 31-AUG-1998; 98US-0098443.

PR 25-JAN-1999; 99US-0117119.

PA (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.  
 (TEXA) UNIV TEXAS A & M SYSTEM.

XX Foster TJ, Hook M, Davis S, Hartford O, Mccrea K, Ni Eidin D;  
 DR WPI; 2000-256637/22.

DR N-PDB; AA293333.

XX Recombinant or synthetic proteins from coagulase-negative staphylococci  
 PT useful for prevention, treatment and diagnosis of staphylococcal  
 PT infections bind soluble and immobilized fibrinogen

PS Claim 2; Figure 2; 104pp; English.

XX Isolated Staphylococcus Sdr cell wall proteins which bind both  
 CC soluble and immobilized fibrinogen are useful for treating or  
 CC preventing coagulase-negative staphylococcal infection such as  
 CC septicemia, osteomyelitis or endocarditis, and for inducing immune  
 CC responses in patients. The cell wall proteins are also useful for  
 CC reducing coagulase-negative staphylococci infection of indwelling  
 CC medical devices such as vascular grafts, vascular stents,  
 CC intravenous catheters, artificial heart valves and cardiac assist  
 CC devices. The cell wall associated proteins are able to inhibit  
 CC staphylococcal adhesion to immobilised extracellular matrix or host  
 CC cells present on the surface of implanted biomaterials.

XX Sequence 1802 AA;

Query Match 16.6%; Score 499; DB 21; Length 1802;

Best Local Similarity 27.1%; Pred. No. 7.2e-19;  
 Matches 168; Conservative 104; Mismatches 275; Indels 72; Gaps 25;

QY 1 SDDEKNDVNNQISITDDNNQIIRKKEETNNYGIKRSDEPSTETNDENEA-TFLQ 59

DB 185 TSTTQODSTKKNPSLNDNLNSSSTTSKESKTDHSTKQAMST-NKSNLDTNDSPTQSE 243

QY 60 KT-PQDNTHLTEEVKRESSVESSNS-----SIDTAQOPSHTTIN-REESVQTSN 108

DB 244 KTSQANNDSSTDNQSAPSKDLSKPSQKVKYKTFNDEPQDVHTTKLTKLPSTSDSS 303

QY 109 VEDSHVDFANSKIKESNTSGKEENTIEQPNKKESTISQPSGYNINDEKI----- 161

DB 304 VMDK--ODYTRSAV---ASLGVDSNETEITNAVRNLDLKAASREICAEIATLAEALK 357

QY 162 --SNODELLNP--INENKARPLSTTSAPSKRTVNOALAE--QGSVNHILKYTD 215

DB 358 DFNMPDYGVDPLALNSQSKNSP--HKSASP--RNLMSLAEPNKGKRVNDRKAITN 412

QY 216 Q--STEGYDDEGVKAKHAENLIYDTPREVDDKVRSGDTMTVDIKNTVPSDLTDSFT 273

DB 413 PTLSLKSNHANNVIMPTSMGEENLKAKEYELDDSIKEGDTFTIKGYQYIRPGLELPAI 472

QY 274 IPKIKDNGSELIAFGTDNKNKQITTFDYVYKYEINAKLKTSLIDSKVNNNTKL 333

DB 473 KTLRSKDGSIIVANGVYDKTNTTTFYFNVDQYQITGSFDLIAIPKRETAIKDNQY 532

QY 334 DVEKKTALSSVNTKIITYEORPENFTANLQSMFTNIDTKNHTVEQTIYINPLASAKET 393

DB 533 PMEYTTINEVYKADFIDYGNKKDNTT---AAYANDVNNKNEVYLLQNNQNRKIA 589

QY 394 NVNISNGDEGSLIIDSTI--IKVYKVDNQLPDS-NRIYDYSEVEDYTNDDYAL- 448

DB 590 KY-----FSYVKNGEFIPGEVKEVETDNTAMVDSFNDLSSNVKDVTSQFAPKVS 641

QY 449 GNNNDVINFGNT---DSPYIKVYSKYDPKNDYTIQCVTVQNTINETGFEFRASY 505

DB 642 ADGTRVDINFRSMANGKKTIVQAVRPTGTGNYT--EYWLTRDGTN--TNFVYRGTK 697

QY 506 DNTIAFSTSSGOGGDLPPERTYKIGDYVEDVDKDIQONTDNEKPLSNVLTLYPDG 565

DB 698 STTYTLNGSSSTAQGDNP---TYSLDGYVMDLKKKNGV--DDEKGLAGVYVTL--KDS 750

QY 566 TSKS---VRTDEGKYQFD 581

DB 751 NNRELQRTTDSGCHYOFD 769

## RESULT 13

AY70119  
 ID AY70119 standard; Protein; 1802 AA.

AY70119;

06-JUN-2000 (first entry)

DE Staph. epidermidis serine-aspartate repeat region protein SdrF.

XX Multicomponent vaccine; immunostimulatory; antibacterial; MSCRAMM;  
 KW microbial surface components recognising adhesive matrix molecules;  
 KW collagen binding protein; CBP; CNA; fibrinogen binding protein;  
 KW Clumping factor A; ClfA; Clumping factor B; ClfB; FnBP;  
 KW fibronectin binding protein; Staphylococcus infection;  
 KW serine-aspartate repeat region protein; SDR protein; SdrF.

XX Staphylococcus epidermidis.

XX Key Location/Qualifiers

FT MISC-difference 12 /note- "Encoded by in-frame stop codon TAA"

FT MISC-difference 28 /note- "Encoded by in-frame stop codon TAA"

FT MISC-difference 1771





```

QY 257 VDIKNTVPSDDLDSFTIPKIKDNGSELIATGTVDNKNKQITVPTDYVDKYENIKAHK 316
Db 240 FKYGQYFRPGSVRLPSOTONLYNAQNIIAKGIYDSTNTTFTNVDQYTNVRSFE 299
QY 317 LTSYIDSKVPNNNTKLDVEKKTALSSVNTITVEYQRPNNRANLQSMFTNIDTKNH 376
Db 300 QVAFAKRNATTDKTAAYKMEVTLGNDYSEELIVDY--GNKKAOPLISTSTNYINNELS 356
QY 377 VEOTIYIN-PLRYSAKETNV-NISGNGDEGSTIIDSTIIKYKVGDNQNL-----PDSN 429
Db 357 RNMATAYNOQPKNTYTKQTFVYNLT-----GYKFNPNAKNEKIEVTDQNOQFVDSFTPDS 411
QY 430 RIYDYSEEDVTNDVYAQLGNNNDVINEGNIDS--PYIIKYI-----SKYDPNKDYTT 482
Db 412 KIKDVTQDFVYI--YSNDNKTATVDLMKGOTSSNKQYIIQVAYAPDNSSTDNGKIDY-- 466
QY 483 IQQTVTQTTINEYTGFEFTASYDNTIAFSTSSGOGGDLPEKTYKIGDYVWEDVDKDG 542
Db 467 ----TLDTKTKYSW---SNSYSNVNGSSTAND-----QKKYNLGDYVWEDVTNKDG 511
QY 543 IONTDNEKPLSNVLYTLTYPDGTS-KSVRTDEDGKYQFDG 582
Db 512 KQDA--NEKGIKGYVYIILKDSNGKELDRITTDENGKYOFTG 550

```

Search completed: March 14, 2003, 12:58:23  
 Job time : 30.0323 secs

Fri Mar 14 14:09:55 2003

us-09-147-405b-13.ra1

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01; Search time 9.3871 Seconds

(without alignments)  
1824.221 Million cell updates/sec

Title: US-09-147-405B-13

Perfect score: 3013  
Sequence: 1 SSDEKNDVNNOSINTD.....PDGTSVTRDEGKYRFDG 582

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents, AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/6D.COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2810.5	93.3	930	US-09-134-001C-5314	Sequence 5314, Ap
2	452	15.0	933	US-08-293-728-2	Sequence 2, Appl1
3	452	15.0	933	US-09-421-868-2	Sequence 2, Appl1
4	349	11.6	345	US-08-856-253-7	Sequence 7, Appl1
5	259.5	8.6	1112	US-08-714-402-2	Sequence 2, Appl1
6	259.5	8.6	1161	US-09-327-536-2	Sequence 2, Appl1
7	220.5	7.3	3686	US-09-134-001C-5080	Sequence 5080, Ap
8	209	6.9	630	US-08-973-462-9	Sequence 9, Appl1
9	202	6.7	676	US-09-134-001C-4318	Sequence 4318, Ap
10	201.5	6.7	1177	US-09-134-001C-5106	Sequence 5106, Ap
11	197	6.5	2314	US-09-268-347-49	Sequence 49, Appl1
12	195	6.5	1335	US-09-134-001C-3716	Sequence 3716, Ap
13	195	6.5	10182	US-09-134-001C-3159	Sequence 3159, Ap
14	192.5	6.4	1786	US-08-973-462-8	Sequence 8, Appl1
15	190.5	6.3	1833	US-08-621-944A-4	Sequence 4, Appl1
16	190.5	6.3	1833	US-08-945-567D-3	Sequence 4, Appl1
17	190.5	6.3	1992	US-08-621-944A-3	Sequence 3, Appl1
18	190.5	6.3	1992	US-08-945-567D-3	Sequence 3, Appl1
19	189.5	6.3	1664	US-09-599-652-2	Sequence 2, Appl1
20	189.5	6.3	1664	US-08-642-846-2	Sequence 2, Appl1
21	189.5	6.3	1664	US-09-264-604-2	Sequence 2, Appl1
22	188.5	6.3	1435	US-08-568-459A-4	Sequence 4, Appl1
23	188.5	6.3	1435	US-08-487-826B-4	Sequence 4, Appl1
24	188.5	6.3	1435	US-09-210-288-4	Sequence 4, Appl1
25	187	6.2	1416	US-09-071-035-404	Sequence 404, App
26	187	6.2	1448	US-09-071-035-402	Sequence 402, App
27	186.5	6.2	3111	US-08-460-309-4	Sequence 4, Appl1

28	186.5	6.2	3111	US-08-125-077-4	Se
29	185	6.1	2048	US-09-268-347-48	Se
30	184.5	6.1	2391	US-08-446-855A-2	Se
31	184.5	6.1	2391	US-09-150-741-2	Se
32	184	6.1	783	5231168-2	Pat
33	177.5	5.9	1087	US-08-264-002-5	Sequence 5, Appl1
34	175.5	5.8	699	US-09-134-001C-4054	Sequence 4054, Ap
35	175.5	5.8	1104	US-09-268-347-28	Sequence 28, Appl
36	175.5	5.8	1104	US-09-268-347-34	Sequence 34, Appl
37	175.5	5.8	1183	US-08-447-031A-2	Sequence 5, Appl1
38	175	5.8	1147	US-08-470-260-5	Sequence 5, Appl1
39	175	5.8	1147	US-08-471-491-5	Sequence 5, Appl1
40	175	5.8	1147	US-08-466-662-5	Sequence 5, Appl1
41	175	5.8	3289	US-08-477-451-2	Sequence 2, Appl1
42	174	5.8	467	US-09-134-001C-3235	Sequence 3235, Ap
43	173.5	5.8	1231	US-08-904-263A-4	Sequence 4, Appl1
44	173.5	5.8	2123	US-08-968-685A-10	Sequence 10, Appl
45	173.5	5.8	2137	US-09-134-001C-4463	Sequence 4463, Ap

# ALIGNMENTS

```
RESULT 1
US-09-134-001C-5314
; Sequence 5314, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCO
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5314
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5314

Query Match      93.3%   Score 2810.5; DB 4; Length 930;
Best Local Similarity 93.6%; Pred. No. 1.3e-158;
Matches 345; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

QY 1 SSDEKNDVNNOSINTDNNQIKKETNNYDGIKRSERESTTNDENAEFLQK 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 SSNEKNDVNNOSINTDDNO-IKKEETNSDAIEKRSKOTIQTNDENAEFLQK 131

QY 61 TPQNTLHTEFEVKESSVESNSSIDTAQPSHTTINRESVTSNDVSHSDRANS 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 132 TPQNTLHTEFEVKESSVESNSSIDTAQPSHTTINRESVTSNDVSHSDRANS 191

QY 121 KIKESNTSGKEENTIEOPNPKVEDSTTSQSGTYNIDEKSNODELNLPINEYENKAR 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 KIKESNTSGKEENTIEOPNPKVEDSTTSQSGTYNIDEKSNODELNLPINEYENKAR 251

QY 181 PLSTTSAPSIKRYTVNOLAEOGSNNHLIKVTDQSTIEGYDSEGVYKAHDAENLIYD 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 PLSTTSAPSIKRYTVNOLAEOGSNNHLIKVTDQSTIEGYDSEGVYKAHDAENLIYD 311

QY 241 VTFEYDVKVSGDMTVDIDKNTVPSDLTDSFTIPKINDNGELIATGYTNKKQTYT 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 312 VTFEYDVKVSGDMTVDIDKNTVPSDLTDSFTIPKINDNGELIATGYTNKKQTYT 371

QY 301 FTDVVDKENTKAHLKLTSLYIDSKSVPPNNKRLDVEYETALSVMKTTVEYQRPENRT 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 372 FTDVVDKENTKAHLKLTSLYIDSKSVPPNNKRLDVEYETALSVMKTTVEYQRPENRT 431
```





RESULT 4  
US-08-856-253-7Sequence 7, Application US/08856253  
Patent No. 6288214

## GENERAL INFORMATION:

APPLICANT: Hook, Magnus  
APPLICANT: Patti, Joseph M.  
APPLICANT: House-Pompeo, Karen  
APPLICANT: Stahanam, Narayana  
APPLICANT: Symersky, Jindrich  
TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS  
TITLE OF INVENTION: AND METHODS OF USE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: U.S.  
ZIP: 77210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/856,253  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/017,678  
FILING DATE: 16-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: TANK:193  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 345 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear

US-08-856-253-7

Query Match 11.6%; Score 349; DB 4; Length 345;  
Best Local Similarity 30.8%; Pred. No. 1.2e-13;

Matches 102; Conservative 61; Mismatches 136; Indels 32; Gaps 15;

QY 213 VDDQ--STEGYDDSGVYKAHDAENLIYDFEVDQKYSQGTMTNVDIDKNTVPDQD 270  
DB 25 ITNQLTNVTVGI-DSGTTYVPHQAGYKLVKGFSPNSAVKGTFTYVPEKELNLGVTS 83  
QY 271 SEITPKIKNSGEIITAGTYDNKNOITTYTDYVDKYEINIKAKLITSIDSKYNNN 330  
DB 84 TAKVPIMAGD-QVLANGYDS-DGNVYTFETDYVNTKDKVAKTLMPAID----PENY 137  
QY 331 TKL-DVEYTAUSS--VKTTIVEYQRPENRPNANTQSMFTNIDTKNHYEQITLYNPL- 386  
DB 138 KKGANTTLATIGSTANTVLDYERKGFYMLSIKGTIDQIDKTNNTYRQTIYVPSG 197  
QY 387 -RYSAKETVNVISNGDEGFTIIDSFTIKVYVGVGNONLPDSNRIYDSEYEDVNDQY 445  
DB 198 DNVIAYVLGNLKPNTDSNALLDQONTSTIKVYVDNADSESYFV-NPNEFDVNSVN 256  
QY 446 AOLGNNDVNIING---NIDSPYIKIVISKYDPN-KDYTTIQQTVMQITINEYTG- 499  
DB 257 ITPFBNQKVEFNTPDDQITTPYIYVNGHIDPNSKGD-----LALNSTLYGVNSNI 309  
QY 500 -FTASYDNTIAFTSTSSGGQG-DLP--PEK 526

DB 310 IIRSMWDENVAFNNGSGSDGIDKRPVPEQ 340

RESULT 5  
US-08-714-402-2Sequence 2, Application US/08714402  
Patent No. 5910441

## GENERAL INFORMATION:

APPLICANT: ROCHA, Claudia  
APPLICANT: FISCHETTI, Vincent A.  
TITLE OF INVENTION: FIBRONECTIN AND FIBRINOGEN BINDING  
TITLE OF INVENTION: PROTEIN FROM GROUP A STREPTOCOCCI  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,402  
FILING DATE: 16-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 016921-097  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1112 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-714-402-2

Query Match 8.6%; Score 259.5; DB 2; Length 1112;  
Best Local Similarity 21.2%; Pred. No. 9.9e-08;  
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;

QY 125 SNTSGKEENTIPQPKVKEDSTTSOPS-----GYNIDKISNODEL--NLPIN 173  
DB 43 ANKGSFEIKKVDQNNKPLPGATSLTSKCKGTSVQSFISNKGIVDAQNLPGTYTLK 102  
QY 174 E-----YENKARPLSTSAQPSIKRYTVN-----QLAAEGS-NVHLLIVTD--OSITE 220  
DB 103 EETAPDGYDTSRTWTYVYENGYTRKLVENPYNGELISKAGSKDVSSLOLEPNKSVWS 162  
QY 221 GYDSDS--GVIAHAENLIYDFEVDQKVS-----GDTMTVIDDKTVSDTLDSF 272  
DB 163 KYKTEVSSCAADPFRNHAAYFKMSFELKCKDKSSETINPDQIVYLDRLNFKGISO- 221  
QY 273 TIKP-IKNSGEIITAGTYDNKNOITTYTDYVDKYEINIKAKLITSIDSKYNNNT 331  
DB 222 -IKRIIYDSNSPLAIGKYAENHQLYFTFDYIAGDKVQLABELSLFLENKEYLEN- 278  
QY 332 KLVETKALSVNKKI-----TVEYQRPNE-----NRTANQSMFTNIDTK 373  
DB 279 -----TSISNFKSTIGGEIITYKGVNVLVYGNSTKESNYITNLSNNGSIESYNE 331  
QY 374 NHTVEQIYVNIPLRYSAKETNVNISNG-----DEGSTIIDSTFIKIVYKVDNQ 423  
DB 332 TGEFVWYVYVNPRTNIPYATMLMGFGRAKRSNTSLENDANTSSAELGELYVEPEGE 391

QY 424 NLPDSNRITYSEEDVT-----NDYAOQLGNNDV-----NINGN--IDSPYIIKYSK 472  
DB 392 KLPSSGYV-----DYTKLRLTRDITAGLNGFOMTKRQIDFGNNGNNAFIKYGK 444  
QY 473 YDPKADYTTIQTIVTMTINEXTGERFAS-----YDNTIAFSTSSGGGCG-- 520  
DB 445 TDOS-----GKPLVQSNLASFSGASEYAFTPVGNGVYFQNEIALSPKSGSGKSE 497  
QY 521 ----DLPEKTYKIGDYVWEDVDKGIONTNDNEKPLSNVLTLYPPDGTSKSVR--TDE 574  
DB 498 FTKPSITIVANKRYAQLRFRKMSIDNV-----PLPEAFELRSSNGNSQKLEASNT 549  
QY 575 DGXYOF 580  
DB 550 QGEVHF 555

SULT 6  
US-09-327-536-2  
; Sequence 2, Application US/09327536  
; Patent No. 6355477  
; GENERAL INFORMATION:  
; APPLICANT: FISCHETTI, Vincent A.  
; APPLICANT: ROCHA, Claudia  
; TITLE OF INVENTION: FIBROECTIN AND FIBRINOGEN BINDING PROTEIN FROM GROUP A  
; FILE REFERENCE: 022927-008  
; CURRENT APPLICATION NUMBER: US/09/327,536  
; CURRENT FILING DATE: 1999-06-08  
; PRIOR APPLICATION NUMBER: US 08/714,402  
; PRIOR FILING DATE: 1996-09-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: SFBP gene  
US-09-327-536-2

Query Match 8.6%; Score 259.5; DB 4; Length 1161;  
Best Local Similarity 21.2%; Pred. No. 1e-07;  
Matches 116; Conservative 102; Mismatches 205; Indels 123; Gaps 23;

QY 125 SNTSGKEENTIEQPNKVKEDSTISOPS-----GYTNIDEKISNODEL--NLPIN 173  
DB 43 ANKGSEIKKVDQNNKPLPGATSSLSKDGKTSVQSFTNDKGIYDAQNLQGTITLK 102  
QY 174 E-----YENKARPLSTTSAOPSIKRYTVN-----QLAAEOGS--NVNHLIKYTD--OSITE 220  
DB 103 EETAPDGDGKTSRTMTVTVYENGCTKLEMPYNGEIIISKAGSKDVSSLSOLENFKMSVVS 162  
QY 221 GYDSDS---GVYKAHDAENLIYDTFEYVDKVS-----GDTMTVDIDKNTVPDLDSEF 272  
DB 163 KYGTETVSSGADYRHHAAFKMSFELKQDKSETINPGDTFVQLDRRLNPNGISGD- 221  
QY 273 TIFP--IKONGEIIATGTYDNKNQOITYTFTDYDKYENIKAHKLTSYIDSKSVPNNT 331  
DB 222 -IPKIIYDSANSPLAIGYHAENHQLIYTFDYIAGLDKVLASLFLFNKKEVLEN- 278  
QY 332 KLDEYETALSSVAKTI-----TVEYORPNE-----NRRANLOSMFTNIDTK 373  
DB 279 -----TSISNFKSTJGGOEITYKGYVNLVYGNESKESNYITNGLSNWGSIESTYTE 331  
QY 374 NHTVEQTIYINPLRYSAKETVNIISNG-----DEGSTIIDSTLIKVKYKVGNDQ 423  
DB 332 TGEVWVYVVPNRNTINPYATNMLMGFRASNTSDLENDANTSALGELQVYEVPEGE 391  
QY 424 NLPDSNRITYSEEDVT-----NDYAOQLGNNDV-----NINGN--IDSPYIIKYSK 472  
DB 392 KLPSSGYV-----DYTKLRLTRDITAGLNGFOMTKRQIDFGNNGNNAFIKYGK 444  
QY 473 YDPKADYTTIQTIVTMTINEXTGERFAS-----YDNTIAFSTSSGGGCG-- 520

DB 445 TDOS-----GKPLVQSNLASFSGASEYAFTPVGNGVYFQNEIALSPKSGSGKSE 497  
QY 521 ----DLPEKTYKIGDYVWEDVDKGIONTNDNEKPLSNVLTLYPPDGTSKSVR--TDE 574  
DB 498 FTKPSITIVANKRYAQLRFRKMSIDNV-----PLPEAFELRSSNGNSQKLEASNT 549  
QY 575 DGXYOF 580  
DB 550 QGEVHF 555

RESULT 7  
US-09-134-001C-5080  
; Sequence 5080, Application US/091344001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5080  
; LENGTH: 3696  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5080

Query Match 7.3%; Score 220.5; DB 4; Length 3696;  
Best Local Similarity 20.3%; Pred. No. 9e-05;  
Matches 134; Conservative 106; Mismatches 241; Indels 180; Gaps 31;

QY 2 SDEEKNDVYNNOSINTDNNQIIKKEETNNYDIEKRSERDTESTTVHDENEATFLQKT 61  
DB 718 ADQCKT-LIRNNHATTEEONEAIRQYEAHSSDAIAKGEAFDTVTVEARNDNGKLIAT 776  
QY 62 PODTHLTREEKSSSVESNSSIDTQAPSHITTIN-REESVOTSDHVEDSHVSDFANS 120  
DB 777 DVNP--TKKARAAYATNSANSKIKIDINNTQATLDERDAIALVNSKD-----E 826  
QY 121 KIKSNTSGKEENTIEQPNKVKEDSTISOPSQGTINIDEKISNODELINLPINEYENKAR 180  
DB 827 AIONINTRAQG-----NDVYTEAQNNG--TNTIQOV----- 854  
QY 181 PLSTTSAOPSIKRYTVNOLAEOGSNVNHLIKYTDOSITEGYDSDSEGVYKAHDAENLIYD 240  
DB 855 PLTVKKNQNAL--ATINAKADEQ-----KRLQANNNAATTE-----EKAADERKNE 899  
QY 241 VTFEVDKVSQGTMTVDIDK--NTVPDLDTSFT--IPKIKONS-----GEIIATGTYDNK 293  
DB 900 AVITANQNTNA--TTNRDVDOAQTTGSGIISAIIPATRIKEDARAAYAKAIAQNOQINS 958  
QY 294 NKQTYFTFDVDVYENKKAH-----LKLTSYIDSK-----VNN 329  
DB 959 NNMA--TTEEKEDALNOVEAHKQALAIITINOAQSTQOVSSEAKNNGINTINODQPAVAKN 1016  
QY 330 NTKLDVEYK-----TALSSVAKTI--VEYORPNEPNTANLOSM 366  
DB 1017 NTKILLEGKNGEKSALIAQTPDATTEEKEQEAHSVAVSAQVATGITHINQANEN----- 1068  
QY 367 FTNIDTKNHTVEQTIYINPLRYSAKETVNIISG-----NDEGSTI 407  
DB 1069 -DVOQELSNAEQIIT-----THTVNVQKRPQARQALIAKTNERKOSAINSDNEGTI 1118  
QY 408 IDDSITIIKVKYVGNQNLPSNRITYSEEDVTYNDYAOQLGNNDVNIIN-----GN 460  
DB 1119 EEKQAKIQ--SLNDAKNLAD--EQITQASQNVNANLNGISNISKIQTQFTKQOARDQ 1175

QY 461 IDSEPIIKVI-----SKYDPNKDDYTTIOQVTMTTINEYGEFRTASTDNTIARST 513  
DB 1176 VNOFQOEKAEKLNSTPHATQOEKODALTRLTQ--AKETALNDINOAGTNOQNDTALTSGI 1233  
QY 514 SSGOG-QODLPEKTYK--IGDYWEDVDKDGIOQNTND--NEKLSNVLTJLTYPDGT 567  
DB 1234 QNIONTOYVNRKQOEAKTINDIYQO--HKOSIONNDATTEKEEVANNLVASOQNVIS 1291  
QY 568 K 568  
DB 1292 K 1292

RESULT 8  
US-08-973-462-9  
Sequence 9, Application US/08973462B  
Patent No. 6191270  
GENERAL INFORMATION:  
APPLICANT: DROITHE, PIERRE  
APPLICANT: DAUBERSIES, PIERRE  
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
FILE REFERENCE: 0660-0125-0 PCT  
CURRENT APPLICATION NUMBER: US/08/973,462B  
CURRENT FILING DATE: 1998-02-06  
EARLIER APPLICATION NUMBER: PCT/FR96/00894  
EARLIER FILING DATE: 1996-06-12  
EARLIER APPLICATION NUMBER: FR 95/07007  
EARLIER FILING DATE: 1995-06-13  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 630  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Polypeptide  
US-08-973-462-9

Query Match 6.9%; Score 209; DB 4; Length 630;

Best Local Similarity 19.5%; Pred. No. 4, 7e-05;  
Matches 116; Conservative 121; Mismatches 206; Indels 152; Gaps 28;

QY 2 SDEKNDVNN-----NOSINTDNNQIKKEETNNYDGIKRSRDRTEST- 47  
DB 59 SDELFELLNSVDVNGEVKENILEESQVNDIDFNSLVKVCQOQHNVKEKVESEVEND 118  
QY 48 -----TNDENE-----ATFLQK-----TPQONTHTLTEEVK 74  
DB 119 EESVEENVEENVEENDGSAVSEESIASVDESIDSIEENVAPTVEIYAPVEIY 178  
QY 75 ESSVSSNSSSIDTAOPSHTTINRESVOTSDNVEDS--HVSDFANSKIKES--NTE 128  
DB 179 ABSVSVASVAVESVAPS-----VEESV--AEVNEESVAENVEIYAPVSESVAE 231  
QY 129 SGRKEENTIEQPNKVKEDSTTSQPSGYTNIDEKIS--NODLLNLPINE-----YENKARPL 182  
DB 232 ESAVNEESVAENVEESVAE-----NVEESVAENVEEYIYAPVSESVAPVEIYAPT 285  
QY 183 STTSAPSTIKRVYNOLAAEOGSNVNHLIKVTQSTIEGIDDSGVYKADAENLIYDVT 242  
DB 286 VESVAPVVEIYAPVSESVAPSV-----EESVAENVEESV-----AEVNEESVA 331  
QY 243 FEVDKAKSG--DTMTVDIKNTVPDLDSTFIPIKIDNGEIIATGTDNKNKOI--- 297  
DB 332 ENVEESVAENVEESVAENVEIYAPV--VEEYIYAPVSESVAEVAVANLMDNLNLG 389  
QY 298 --TYTFDYV--DREYENIKAKHLKLTYSIDKSKVPNNNTKLDVEYKTAHSV----- 344  
DB 390 IETFEIKDSILNEIEVEKENV--VTTLIEKVE-----ETTAESVTFSTFNLIEIOENTIT 442  
QY 345 NKITTYEYQRPNENKRTANLOSMFTNIDTKNHTVEQTIYINLKRSAKETVNVNISGNDEG 404

DB 443 NDTIEKLEELHEN---VLSAALENTOSEEKEVYDIEVEKEEVAATLIEVEQAESE 499  
QY 405 STIIDDSITIKY-----KQVDN-QNLEPDS--NRKYDSEY-----EDVYND 443  
DB 500 S-----ESTITTEFLENNAVESNEKVAENLEKLETFNVLVDVEETVEISGLENN 555  
QY 444 D-----YAOQGN-----NDVINFGNDSPYIIVKISKYDPNKDDYTTIOQTV 487  
DB 556 EMDKAFSEIFDPNVKGIQENLLTGMRFSIETSIIVIOSEKVDLNNVSSIINDI 610

RESULT 9  
US-09-134-001C-4318  
Sequence 4318, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4318  
LENGTH: 676  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4318

Query Match 6.7%; Score 202; DB 4; Length 676;

Best Local Similarity 22.1%; Pred. No. 0, 0.00013;  
Matches 116; Conservative 82; Mismatches 191; Indels 136; Gaps 23;

QY 2 SDEKNDVNNNOSINTDNNQIKKEETNNYDGIKRSRDRTESTTNDENATFLQKT 61  
DB 77 SFEERLDKVDNOSIDRIIND-----AKDKNNH--LKSTDSASVSTEDDT-----S 123  
QY 62 POONTHTLTEEVESVSSNSSSIDTAOPSHTTINRESVOTSDNVEDSHVSDPANSK 121  
DB 124 EKNDNMTKDLKILDLSDIAKVNDRQOGENSASRPSDS--TTIEKDS-----NNK 175  
QY 122 IKESN-----TESGREENTIEQPNKVKEDSTTSQPSGYTNIDEKISNODLLNLPINE 175  
DB 176 VHDITNASTRNATTDSEESVYDKLKDIOQFKSDS-----NNKLSQSDQOQASPSNKN 228  
QY 176 ENKARPLSTISAPST-----KRVYNOLAAE-----OGSNVNHLLIKVTD-----Q 216  
DB 229 ENKKEESTTNNOSDSKDKSDKNSDGRSLTERIASDTDOIRSKDOH---VIDEKODIO 285  
QY 217 SIT---EGYDSEGVYKADAENLIYDVT-----EYDDKVKSGDTMTVD----- 258  
DB 286 AITRSLOGSKIEKALAKVSDNQPDSNINNNKMLNRLSDIKVEDNNTLSDDKQALK 345  
QY 259 --IDKNTVPDLDSTFIPIKIDNGEIIAT-----GTYDNKNKO 296  
DB 346 QEIDKTKOSIDROENIILIDOLNGASNKQAKTEIDILNSVFSKNEVEDIMKRIKTGRSNE 405  
QY 297 ITTYTFDYVQYENIKAKHLKLTYSIDSKYPNN-----NFKLDVEYKTAHSV-- 343  
DB 406 IANOIAKOIDGLALTSDDILKSLMDSKOKESLIKOLLTRIGANDADRIAKLLSONL 465  
QY 344 VNKITTYEYQRP--NENRTANLOSMTNI--DTKNHTVEQTIYINLPRYSAKETVNVNISGN 400  
DB 466 SNSQIVOLKRRHFNFSOGTARADDLINGVINDAKD-----KRAIETTLQTRIN 513  
QY 401 GDEGSTIIDSTIIKVKYKVDGNQNLPSNRITYD--YSEYEDVYND 443  
DB 514 KDKAKIAD--VIARVOK-----DKSDIMDLHSAIEGKAND 548

RESULT 10  
US-09-134-001C-5106  
Sequence 5106, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIORITY APPLICATION NUMBER: US 60/064,964  
PRIORITY FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5106  
LENGTH: 1177  
TYPE: PRF  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5106

Query Match 6.7%; Score 201.5; DB 4; Length 1177;  
Best Local Similarity 19.1%; Pred. No. 0.00029;  
Matches 131; Conservative 121; Mismatches 242; Indels 191; Gaps 30;

QY 7 NDVNNNSITDNNQIIRKEETNNYDGIKRESEDTSTTNDENATFLQKT----- 61  
DB 134 NNNNNNSVFNQNDNKKSSQQRKSTIOSENIRKANKTKTSTST---PEFTYLNHSFKSE 189  
QY 62 -----PODNHILTEEVKESSESVSSSIDTAQOPSHITINREES 102  
DB 190 VPSAIFGTRKRPPIENGVIPEHEKELNDKEIVQODEVSHSTKSIDASKVNSNDNNIEK 249  
QY 103 VQ-----TSDNVEDSHVDFANSKIKESNTSGKEENTT----- 136  
DB 250 NQAKQOQTAAOTESSEMHVVEKSNVQTKRKTPN--YSKVNTNINENIYASQIVEEI 307  
QY 137 -----EOPNKVYKEDST-----TSOPSGYT--NID-EKIS 162  
DB 308 RREBERKVLQRRFKKALQOKRQONQOSEDSIOKAIDEMAKQAHYTGESSIDLENES 367  
QY 163 NODDELLNPINEYENKARPLSTTSNOPSIRKVTYNOGLAEQSNVNHILKTDOSTEGY 222  
DB 368 NODSSSN--SLEKQSNSSNIDNKEAQNTPLENEIDLDITSDV--YKVNTEE-TESEK 421  
QY 223 DDESGVIAH-----DAENLIYDVF--EVDKVKSGDTMTVD--IDKNTVPSDLTDSFTI 274  
DB 422 NDEDLVSSNHTHSNDAAE--VEDAEYHLLDNRQONQSNQSDDIISKSSSTSNMIDNAIS 479  
QY 275 PKIDNNGEIIATGTYNKKNQITV-----TFDYVDKYENIKAKHLKL 318  
DB 480 ASVNNMTER--AKSNEDKNQTEITHLDGTSAKVSEKIESENENHLLBODKNV--LKNV 535  
QY 319 SYIDSKVPPNNNTK-----LDVEYKALSSVVKKTIVVEQRPMENTALQ 364  
DB 536 NSLWSSNDGTQTRQRFGRGSRPENVLTPSDKRRMDQNRKQVSPLEKPEKQANNAHR 595  
QY 365 SMFTNIDRKNHTVEQTIYINPLKYSAKETNPN--ISNGNDEGSTIIDSTIIKVVYVGNQ 423  
DB 596 K-----DESKKSEEFKIN--TNRETDSNYSNGNIEHD--MNSSDEHVIYETPKQ 644  
QY 424 NLPDSNRILYDSEYEDVTNDYDQALGNNDVNIENGINDSPIYIKVISKYDPNKKDYTTI 483  
DB 645 ODEQLOKQDDFHFE--NANHAKINNSNET--GN-----QENI 678  
QY 484 QQTVMQTTINEYGEFTATYDNTIAFSTSSGQGGDLPEPKTKIDY---VWEVD 539  
DB 679 SHSKRSQYSTRE-----SKNIDTQTSNSTSNQFORIRKGNIKLPSTYOLLEAPEPH 732  
QY 540 KDGIQNTDNNEK--PLSNVLVTLTYP 563

DB 733 KD--QDWIDNKKQELNDALYYFNVP 755

RESULT 11  
US-09-268-347-49  
Sequence 49, Application US/09268347  
Patent No. 635182  
GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena M.  
TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
FILE REFERENCE: 1038-860  
CURRENT APPLICATION NUMBER: US/09/268,347  
CURRENT FILING DATE: 1999-03-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 49  
LENGTH: 2314  
TYPE: PRF  
ORGANISM: Haemophilus influenzae  
US-09-268-347-49

Query Match 6.5%; Score 197; DB 4; Length 2314;  
Best Local Similarity 21.1%; Pred. No. 0.0012;  
Matches 135; Conservative 101; Mismatches 261; Indels 144; Gaps 30;

QY 4 EEKNDVINNNOSIN--TDNNQIIRKEETNNYDGIKRESEDTSTTNDENATFLQKT 62  
DB 379 ETOADKLTDDNNIGVNTNNGLKVKLAKNLGLSGL-----TVSTKNTLASE---KVTY 428  
QY 63 QDNHILTEEVKESSESVSSSIDTAQOPSHITINREESVQTSN---VEDSHVSDFA 118  
DB 429 GSGNNTALQGGGLFTFTTNASTD-----KTYVGDGLKFTNNTALEDT--TRIT 479  
QY 119 NSKIKESNTSGKEENTIEQP---NKVKESTISQPSGYTNIDKISNODDELLNPLINE 174  
DB 480 KDKIIFSNAKATVDEN---KPYLDKDKLVGNSTLNNGLT--VNNITGSKNQIOV--- 531  
QY 175 YENKARPLSTTSNOPSIRKVTYNOGLAEQSNVNHILKTDOSTE---GYDSSGVIAK 231  
DB 532 -----GADGIFADVN-----VWVNAAFGTTTITEEIGFADAGKVDK 572  
QY 232 H-----DAENLIYDVFVDDKVKSGDTMTVDIDKNTVPSDLTDSFTIPKIR---DMSG 282  
DB 573 KSPYIDKQLOVGVYKTKRDSGINAGDOKISNVKAT---DQDAVYTKQLOKVOODADG 629  
QY 283 EIIATGTYNKKNQITV-----YTFDYVDKYENI---KAHLKLTYSIDSKVPPNNTKLD 334  
DB 630 ALQSFIDEKQGEFTISNLVSNNGTPTMFTETITFAGEGSIISNDIAKGV-----KVG 684  
QY 335 VEKTAALSSVVKKTIVVEQRPMENTALQSFHTIDRKNHTVEQTIYINPLKYSAKETN 394  
DB 665 IDPGLTLTPKLTV-----GSDKDGKTQVLVEQVASGNGKNTIRGVSPLPSI--TN 735  
QY 395 VNISNGDEGSTIIDSTIIKVVYKVGDNQNT-----PDSNRILYDSEYEDVNDYDQALG 449  
DB 736 AGCVATTQGNITITDEKSKRAASIGDILNTGFNLKNNNSNGVFTYTVTFID-----G 791  
QY 450 NNDVNIENGINDSPIYIKVISKYDPNKKDYTT-----IQQVTVMQTTINEY 496  
DB 792 NATTAKVYDEFNQSKV---TYDVANDEKTIETLGDNGKTNKIGVKTTLTTNNANGK 847  
QY 497 TGEFTATYD-----NTAFSTSSGQGGDLPEPKTYAI--GDYVED---VD 539  
DB 848 ATNFSTNDALVNAKDAIENLNTLAKELHTTKGTADALQ--TFKVKKDGATVDEFTIVG 906  
QY 540 KDGIQNTDNNEKPLSNVLVTLTYPDGTSGSVRTDDEGKYOF 580  
DB 907 KDGTQN-----GKTYNTLKLKENGILTVATINNDGIVTF 939

RESULT 12  
US-09-134-001C-3716  
Sequence 3716, Application US/09134001C

Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3716  
LENGTH: 1335  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3716

Query Match 6.5%, Score 195; DB 4; Length 1335;  
Best Local Similarity 20.8%; Pred. No. 0.00081;  
Matches 160; Conservative 97; Mismatches 254; Indels 260; Gaps 43;

7 NDVYNN--NOSINTDDNNO-----IHK--KEETNNYDGIKESDERTSESTVNDENAEAF 57  
100 NEISSNOQOUSTDDANONOTNVTKNQOQETN-----DLTQEKSTSTIDNLODETOSV 154  
58 LOKTPOD---NTHLFEERKESSESVSSNSIDTAQOPSHTTINREESVOTSDNVEDSHV 114  
155 AKENKDKGANNANNOQOKMTASQPSNOAIET-----QTNNDNESQOKSOQV-TSPQ 208  
115 SDFANSKIKESNT-----ESGKEENTIEQPNKVKEDSTTSPSGY----- 154  
209 NETAPKSVNTWASGYNFDYDEDDSDSTHLEPISLNNVNTSKOTSTYKKEPAQRYT 268  
155 TNI--DEKISNDELINLPINYEKKARPLSTTSAPSKRTYVNOALAEQSNVHLK 212  
269 TMTVKETASNOATIDTQTFPSATAPRYVSVS-SQKTSLSLPYTPKRVSSINNYR 327  
213 VTDQS---ITEGYD-----SEGYIKAHDAENILYDVFEEVDKVKSGDTWT 256  
328 KKNMKAPEEEDYTSYFPRKYGRNGVGRGIV-VHDTAN-----DNST 370  
257 VD-----IDKN-----TVPSDLT-----DS 271  
371 IDGEIAFMKRYTNAFVHAFVDGNRIETAPTDYLSMGAGYGNQRFINVEIYHTHDYS 430  
272 FTIPRIKNSGEIAT-----CTYDNKKQIYTF-----TDYDKYENIKA 313  
431 FA--RSMNNYADYAATQLOLYNLIKPDSEADNGGTWTAHAIISNLTGTDHADPHOYLS 488  
314 H-----LKITSYIDSK--VP--NNNTKLIDVEYKTAASSVNTITVEYOR----- 354  
489 HMYSAELDYLDIYEKRLITKQVAPMGSTSTKPSQSPKPS--GGTNKKLFPVSARGAQIK 547  
355 PNEENTANLOSMTNI-DTKNH---TVEQTIYI-----NPKRIYAKETN----- 354  
548 PTNN-----GLYTYTYDDGKHTQDQOKTLSTVKTATLGNKFYLVEDVNSKKYGMWK 601  
395 -----VNISNGDEGSTI-----IDSTIIKVKYVDNOMLPSNRLYD 454  
602 QGDVYVNTKAPYKVAQOTINVAAGSTLYVPKQTPKQVASKVSGTG-----NOTFKA 653  
435 SEVEDYTNDDYAOAGNNNDVNIINFGNIDSPYIIKIVISKYDPKDDYTTIOQVYTMOT--- 491  
654 TROQOQIDKATFLY---GVVNGSGMISKYLLTTLTASKPSMPTKPS--TNNQLTVTNNSGVA 708  
492 TINETYGERRTASD-----NTI--AFSTSSGQGGDLPEPEKTYIGD-----YWED 537  
709 QINAKNSGLYTYVYDTRKGTQNTQIORTLSVTKAATLGD---KKFYLVGDNVTGTNYGMWK 765  
538 VPKDQONTNDNKKPLSNVLYTLTY-----PDGSKSVRDEDECK 577

Db 766 QD-EVIYNT-----AKSPKINQTYNVKRGVKLHTVPMGTYNQVAGTVSGK 810

RESULT 13  
US-09-134-001C-3159  
Sequence 3159, Application US/09134001C  
Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCO

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3159

LENGTH: 10182

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3159

Query Match 6.5%, Score 195; DB 4; Length 10182;  
Best Local Similarity 19.8%; Pred. No. 0.01;  
Matches 133; Conservative 102; Mismatches 257; Indels 180; Gaps 26;

2 SDEKNNDVYNNOSINTDDNNOIHKKEETNNYDGIKESDERTSESTVNDENAEATLOKT 61  
5081 SONE--SVTNESYINIAEPEKQHAFTALNNAKEIYNEDQATIDANSIOKAQALITTKN 5138  
62 PODNTHLFEERKESSESVSSNSIDTAQOPSHTTINREESVOTSDNVEDSHVSDPANSK 121  
5139 ALDG---EQQLRAK--ENAOEITLNLQTDQAGNSKGLINSQOTREVASOLAKAK 5192  
122 IKESNTSGKEENTIEQPNKVKEDSTTSPSGYTNIDEKISNDELINLPIN--YEN- 177  
5193 -----ELKKVME-----OLNHLINKNOMINSKRFINEDANOQAYSNA 5231  
178 --KARPLSTTSAPSKRTYVNOALAEQSNVHLKVTDOSTIEGYDSEGYIKAHDAE 235  
5232 IASAEALKKNSQNPBLDKQVTEBOAINNINSAINL--NGEAKLTAKKEDVAST----- 5283  
236 NLIYDTFEYDDKVKSGDTWTVDIDKNTVPSDLTDS-----FTIPRIKNSGEIATG 288  
5284 NNLSGLTNE--QKTKENQAVNGAQTRODYANKRLDAEALDOQMOTLRDLVNNQNAIHS 5341  
289 TYDNKKQIYTFYDVKYENIKAHKLKLTSYIDKSKVPNNNTKLVEYKTAAL--SSVK 346  
5342 NYENEDSTOKNTYDNAID-----NGSYTYTGQHNPELKNKSTIDQ 5380  
347 TIVERQRPN--ENRTANLOSMTNIDTKHTEQTIYINPLYSKAKETVNTISG----- 399  
5381 TIS-----RINTAKNDLHGVEKLRDQTAQOELGOLGYLNDPKSGSESLVNSNTRSEV 5436  
400 -----NGDEGSTIIDSTIIKVKYVDNOMLPSNRI--- 431  
5437 EEHLNBAKSLNNAMKOLRDKVAKETVVKQSSDYINDST---EHQRGVQALOEAENITNE 5493  
432 ---YDSEYE---DYTHDDVYAOAG-----NNNDVNI-----FGNI 461  
5494 IGNPITLNKSEIEKLOOLTQAOANALOGSHLEAKKNATITGINKLALANDAQOKAIENV 5553  
462 DSPYIIKIVISKYDPKDDYTTIOOT---VTMOTTINETYGERP---TASYDNTIASF 512  
5554 QAOOTIPAVNOQLTLDBREITAMQALBRKYGOQNNVHQSNYTRNEDQPKHNTDN----- 5609  
513 TSSGQGGDLPEPEKTYIGDYVEDVDK--DGIONTNDNKKPLSNVLYTLTYPDGTSKSV 570  
5610 VQAGQ-----TIIDKLQDPINNKKEIHOAINQINTQTALSGEKKLH 5651

OY 571 RTDEGKYOFG 582  
DB 5652 TDEESTNRQIEG 5663

## RESULT 14

US-08-973-462-8  
Sequence 8, Application US/08973462B  
Patent No. 6191270  
GENERAL INFORMATION:  
APPLICANT: DROULHE, PIERRE  
APPLICANT: DAUBERSTES, PIERRE  
TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
FILE REFERENCE: 0660-0125-0 PCT  
CURRENT APPLICATION NUMBER: US/08/973,462B  
EARLIER FILING DATE: 1998-02-06  
EARLIER APPLICATION NUMBER: PCT/FR96/00894  
EARLIER FILING DATE: 1996-06-12  
EARLIER APPLICATION NUMBER: FR 95/07007  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1786  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Polypeptide  
US-08-973-462-8

Query Match 6.4%; Score 192.5; DB 4; Length 1786;

Best Local Similarity 18.9%; Pred. No. 0.0016;  
Matches 128; Conservative 138; Mismatches 261; Indels 149; Gaps 26;

OY 14 OSINTDNNQIKKEETNNYDGIERSDRTSTNVDENEATFLOKTPQDNTHTLDEEV 73  
DB 549 ESVEENVEESVAENVEESVAENVEESVAENVEE--SVANVEEIVAPVEEIVAPTEVEI 606  
OY 74 KESSVESSESSIDTAQPSHTTIRRESVOTSDNVEDSHVDFPNSKIKESNTESGKE- 132  
DB 607 VAPSVESVAPSVESVVEE--NVEESV--AENVEES--VAENVEESVAENVEEVI 654  
OY 133 -----ENTIEO--PNKYKEDSTSGPSGTINIDKIS--NODELINLFINYEKA 179  
DB 655 VAPVEEIVAPVEEIVAPSVESVAPSVESVVEEENVEESVAENVEE--SVANVEESVA 712  
OY 180 RPLSTTSAO-----PSIKRV--TVNQLAEOGSNNHLIKYVDQSTTEGVDSEGYI 229  
DB 713 ENVEESVAENVEEIVAPVEEIVAPTEVEIAP--SVESVAPSVESVVEEENVEESV-- 767  
OY 230 KAHDAENLIYDTFEVDKVKSG--DTMIVDIDKNTVPSDLTSTTIKIKNSGEIAT 287  
DB 768 ---AENVEESVAENVEESVAENVEESVAPVEEIVAPS--VEESVAPSVESVAPVAT 821  
OY 288 GTYDKNNKOI-----TYFTFDYV--DKYENIKAKLKTYSIDKSKYNNNTKLDVYKTA 341  
DB 822 NISDULSLNLOGITEETEEKSIDLNEIEVKEVNTT-----ILENVEETTAESVTF 874  
OY 342 SSV-----NKTIVEXORPNEN--RTANLQSMFTNIDTKN----- 374  
DB 875 SNILBEIQENTITNDTIEKLELHENVLSALENTQSEEEKKEVIDVIEEVEKAVATTL 934  
OY 935 -HTVQO-----TIINPLRISAKETNNVNSGNGDEGSIITIDSTTIKYYKGDNON 424  
DB 935 IETVQAEKRSANTITEIFENLEENAVESNENVAENLEKMETVNTVADKVEETEVEISG 994  
OY 425 LPSDRIDYSYEDVTDNDVYQNLGNNNOVNNFGNIDSPYIIKYSKYDPKKDYYIT- 483  
DB 995 ESELENNEDKAFSEIF--DNVKGIOENLIGMFRSILETSIYISSEKVDLENVVSIL 1052  
OY 484 -----OQTVMTQITINETYGEFTASVDNTIAFTSSG----- 516  
DB 1053 DNIEMKGLINKLENISSTBEVQETVTEHVEQNNYVVDVPAKMDQFLGILNEAGLKE 1112

OY 517 -----QGGDLPEPEKTYIGDIYWEVDUKDGIQNTDNKEPLSNVL-----VITL 561  
DB 1113 MEFNLEDVFKSSSDVITVEEIK--DEPVQKEVEKEVTSIIEMEENIVDLSEEEKEDLTK 1171  
OY 562 YPDGTSKSVRTDEGK 577  
DB 1172 MIDAVEESIEISSDSK 1187

## RESULT 15

US-08-621-944A-4  
Sequence 4, Application US/08621944A  
Patent No. 6440425  
GENERAL INFORMATION:  
APPLICANT: SASAKI, Ken  
APPLICANT: HARKNESS, Robin E.  
APPLICANT: LOOSMORE, Sheena M.  
APPLICANT: KLEIN, Michel H.  
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER  
MEMBRANE PROTEIN OF MORAXELLA  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sam & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/621,944A  
FILING DATE: 26-MAR-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/478,370  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-587  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO. 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1833 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-621-944A-4

Query Match 6.3%; Score 190.5; DB 4; Length 1833;

Best Local Similarity 19.7%; Pred. No. 0.0022;  
Matches 135; Conservative 114; Mismatches 254; Indels 181; Gaps 34;

OY 3 DEEKDVYNNOSINT-----DDNNQIKKEETNNYDGIERSDRTSTNVDENEAT- 56  
DB 602 DKDKSNASINDILNTGKLNKNNNPDI--DEVSTYDIYDFANGNATYATYHDPANNTS 658  
OY 57 -FLOKTPQDNT--HLTEEVESSESSVSSSIDTAQPSHTTIRRESVOTSDN---VE 110  
DB 659 KYVIDVNVDDTTHLGTGDNKKLGKTKTKKTSANGTATNEN---VNSSDDALVN 714  
OY 111 DSHVSDPANSKIKESNTSGKEENTIE--QPKVKVEDSTTSQSGVTNIDEKISNQ----- 164  
DB 715 AKDIAENMLTAKELHTIKGTADTALQTFYAKKVDENNADANAIYQKANNQNVMTL 774

QY 165 ----DELINLPINEYENKARLSTTSAP-----SIKRYVN--OLA----- 202  
 Db 775 TLKGENGLINTDKNNGVTEGINTSGIKAKSTLNDGGLSIKNFTGSEQIOVGADGVKF 834  
 QY 203 -----OGSNVHLIKVTDOSITEGYDSEGV1----- 229  
 Db 835 AKVNNNGVVGAGIDGTIRITRDEI--GFTGNGSLDKSRPHLSKDGINAGKKITINIOG 892  
 QY 230 ---KAHDA--ENLIYDVFEDDKVS-GDTMTVDIDKNTVPSDLNDSFTIPKIKONG 282  
 Db 893 FIAONSHDAVYTGKIVYDLKTELENKISSTAKAONSLHEFSVADGONNFTVSN----- 946  
 QY 283 EITATGYDNKKNKOITTFYDVYDKYENIKAKLITSYDKS--KVPNNNTK-LDVEYKT 339  
 Db 947 ---PYSSYDTSKSDVITFFAG-----EN-----GITTKVKNKGVYRVGIDQTKGLTTPKLT 993  
 QY 340 ALSSVNTTIVEYORPENRANLOSMTNIDTKNHTVEOTIYINPLRYSAKETVNIISG 399  
 Db 994 VGNNGKGIYIDSON-GONTITGLSNTLANVTNDKGSVRT----- 1033  
 QY 400 NGDEGSTIIDD-----STIIKVKVGNONLPSNRIVDY-SEYEDVYNDYAOQGNND 453  
 Db 1034 --EQGNITIKDEDKTRASIVLVLSAG--FNLOGNEAVDFVSTYDTVNFAD---GNATT 1085  
 QY 454 VNINFGNIDSPYIIKIVISKYDPNKKDYTIO-----QYVMTQT-----INEX 496  
 Db 1086 AKVYTD--DTSKTSKV--YDVNVDD--TTIEVKDKKLGKFTYTLTSTGTGANKFALSNOA 1140  
 QY 497 TGEFRTASVD-----NTIAFSTSSGOGGDLPEKTYKIGD--YVMEVDVKDGIONTND 548  
 Db 1141 TGDALVKASDIYVAHLNLTSGDIOFPAKASQANNSAGYVADGKNVYIDSTDNKYIOAKND 1200  
 QY 549 -----NEKPLSNVLVTLVPPDGT 566  
 Db 1201 GTVDKTEVAKDKLVAAQOTPDGT 1224

Search completed: March 14, 2003, 13:00:50  
 Job time : 24.3871 secs

GenCore version 5.1.4-p5-A578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 13:00:30 ; Search time 8.60176 Seconds  
(without alignments)  
3086.325 Million cell updates/sec

Title: US-09-147-405b-13

Perfect score: 3013  
Sequence: 1 SSDEKNDVNNNOSINTD.....PDGTSKSVTDEDKYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCT05\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	576	19.1	1349	10	US-09-815-242-5898
2	576	19.1	1349	10	US-09-815-242-13137
3	499.5	16.6	932	10	US-09-815-242-5578
4	499.5	16.6	932	10	US-09-815-242-12438
5	455	15.1	841	10	US-09-815-242-5779
6	455	15.1	841	10	US-09-815-242-17751
7	449	14.9	1021	10	US-09-815-242-5471
8	449	14.9	1021	10	US-09-815-242-12544
9	376	12.5	767	10	US-09-815-242-5899
10	376	12.5	767	10	US-09-815-242-13140
11	349	11.6	345	10	US-09-815-820-7
12	319.5	10.6	978	10	US-09-815-242-3456
13	319.5	10.6	1001	10	US-09-815-242-12686
14	308.5	10.2	1018	10	US-09-815-242-5797
15	308.5	10.2	1018	10	US-09-815-242-12838
16	209	6.9	630	9	US-09-742-096-5
17	206.5	6.9	2368	10	US-09-815-242-5635
18	206.5	6.9	2368	10	US-09-815-242-12389
19	201	6.7	2478	10	US-09-815-242-5816

20	201	6.7	2478	10	US-09-815-242-12967	Sequence 12967, A
21	199	6.6	665	9	US-09-820-843A-107	Sequence 107, App
22	198.5	6.6	807	9	US-09-820-843A-108	Sequence 108, App
23	196.5	6.5	1139	9	US-09-820-843A-15	Sequence 15, App1
24	194	6.4	2150	9	US-10-135-322-17	Sequence 17, App1
25	192.5	6.4	1786	9	US-09-742-096-3	Sequence 3, App1
26	188.5	6.3	1435	9	US-10-153-273-4	Sequence 4, App1
27	187.5	6.2	2344	10	US-09-815-242-12713	Sequence 12713, A
28	186.5	6.2	1111	10	US-09-815-242-12955	Sequence 12955, A
29	186	6.2	836	9	US-09-820-843A-109	Sequence 109, App1
30	186	6.2	861	9	US-09-858-525A-2	Sequence 2, App1
31	186	6.2	871	9	US-09-858-525A-1	Sequence 1, App1
32	182.5	6.1	1421	10	US-09-924-154-13	Sequence 13, App1
33	180	6.0	789	9	US-09-995-587A-1	Sequence 45, App1
34	175.5	5.8	1183	9	US-09-870-759-45	Sequence 45, App1
35	175	5.8	1338	10	US-09-402-100-4	Sequence 4, App1
36	174.5	5.8	2437	10	US-09-815-242-5834	Sequence 5834, Ap
37	174.5	5.8	6281	10	US-09-815-242-12996	Sequence 12996, Ap
38	173.5	5.8	837	10	US-09-815-242-5883	Sequence 5883, Ap
39	173.5	5.8	875	10	US-09-815-242-13080	Sequence 13080, A
40	173.5	5.8	2122	9	US-09-813-214A-9	Sequence 9, App1
41	173	5.7	2076	10	US-09-815-242-5815	Sequence 5815, Ap
42	173	5.7	2186	10	US-09-815-242-12913	Sequence 12913, A
43	171	5.7	2025	10	US-09-815-242-5703	Sequence 5703, Ap
44	171	5.7	3158	10	US-09-815-242-12611	Sequence 12611, A
45	171	5.7	5795	10	US-09-815-242-12610	Sequence 12610, A

## ALIGNMENTS

RESULT 1  
US-09-815-242-5898  
Sequence 5898, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trewick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA 011A  
CURRENT APPLICATION NUMBER: US/09/815, 242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191, 078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206, 848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207, 727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242, 578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253, 625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257, 931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269, 308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 5898  
LENGTH: 1349  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5898

Query Match 19.1%; Score 576; DB 10; Length 1349;  
Best Local Similarity 30.1%; Pred. No. 9; 8e-24;





```

APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5578
LENGTH: 932
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5578

```

```

Query Match      16.6% Score 499.5; DB 10; Length 932;
Best Local Similarity 27.7%; Pred. No. 8.6e-20;
Matches 160; Conservative 95; Mismatches 212; Indels 111; Gaps 24;

```

```

OY 35 GIEKSEDRTESTNVDEENATFLQKTPDNDHLEEEVKESSVSNSSIDTAQPSH 94
DB 45 GHEAKAAEHTNGELNOSKNETT-----APSEN-----KTEKVDSRQ--- 82
OY 95 TTINREESVOTSDNVEDSHVSPFANSKIKESNTESGKEENTIEOPKVKVEDSTTSOPSGY 154
DB 83 -----NNIEOSTTSD--QPKVNESDNTSVKE--TTEEP---QNTSTQPTKQ 122
OY 155 TNIDEKISNDELNLPINEYENKARPLSTSAQPSIK-----RYVNOLOAA-EOGSNNY 208
DB 123 NN--DAMAKD--NLAONISTQADKDVSTTPKTTIKPRTLNRMVNVNVAAPQOQTNN 177
OY 209 HLIKVTDOSI-----TEGYDSEGVYKAHDAENLIYDVFVEVDKVKSGDTM 255
DB 178 DKVHFTNIDIALDKGHVNTGTGTEFWATSSDYLK-----LKANYTIDDSVKGDTF 229
OY 256 TVDIDKNVPSDLTDSFTIPKIKDNGSEIATGTYNKKNKQIYTTTDDYDKENIKAKHL 315
DB 230 TFRYGGYFRPGSVRLPSOTONLYNAOGNIIAKGIYDSESTTYYTFTFNVDQYTNISGSE 289
OY 316 KILSYIDKSKVPPNNNTKLDVEYKATLSSVKNKTITVEYQRENERTNLQSMFNTIDTKNH 375
DB 290 EOYAFKREKRENTATDKATYPRVEVTLGNDKYSKNYIVD---GNOKGOQLSSTNYINNEDL 346
OY 376 TVEQTYIYN-PLRYSAKETNV-NISGNGDEGSIIDDSIIKVKYKVDNON-----LPD 427
DB 347 SRMNTYVYVNOPKRTYKKEFTVNTLT-----GYKFNDAKFKIYEV-TNQNQVDSFTPD 400
OY 428 SNRIYDYSEEDYTDNDYQALGNNDVYNIFG--NIDSPYIIVISKYIDPNKDDYTTIQ 485
DB 401 TSKLTVDVTKFKIT--YSNDNKTATYVDLNLGQSSSDKYITIOQVAVPDNSSTDNKGIDY 457
OY 486 TYVMQTTINEYGEFTASTDNTIARSTSSGQGOGLDPEKTKIKIDYVWEDVDKXION 545
DB 458 TLETQNGKSSWSN-----SYSNVNGSSTANGD-----OKKYNLDYVWEDVTNKGKOD 505
OY 546 TNDNEKPLSNVLTLYTPDQGS-KSVRTDEGDKYQFQD 582
DB 506 A--NENKIGKGVYIILKDSNKGELDRITTTDENGKIOFTG 541

```

## RESULT 4

```

US-09-815-242-12438
Sequence 12438, Application US/09815242
Patent No. US20020061569A1

```

## GENERAL INFORMATION:

```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 12438
LENGTH: 932
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12438

```

```

Query Match      16.6% Score 499.5; DB 10; Length 932;
Best Local Similarity 27.7%; Pred. No. 8.6e-20;
Matches 160; Conservative 95; Mismatches 212; Indels 111; Gaps 24;

```

```

OY 35 GIEKSEDRTESTNVDEENATFLQKTPDNDHLEEEVKESSVSNSSIDTAQPSH 94
DB 45 GHEAKAAEHTNGELNOSKNETT-----APSEN-----KTEKVDSRQ--- 82
OY 95 TTINREESVOTSDNVEDSHVSPFANSKIKESNTESGKEENTIEOPKVKVEDSTTSOPSGY 154
DB 83 -----NNIEOSTTSD--QPKVNESDNTSVKE--TTEEP---QNTSTQPTKQ 122
OY 155 TNIDEKISNDELNLPINEYENKARPLSTSAQPSIK-----RYVNOLOAA-EOGSNNY 208
DB 123 NN--DAMAKD--NLAONISTQADKDVSTTPKTTIKPRTLNRMVNVNVAAPQOQTNN 177
OY 209 HLIKVTDOSI-----TEGYDSEGVYKAHDAENLIYDVFVEVDKVKSGDTM 255
DB 178 DKVHFTNIDIALDKGHVNTGTGTEFWATSSDYLK-----LKANYTIDDSVKGDTF 229
OY 256 TVDIDKNVPSDLTDSFTIPKIKDNGSEIATGTYNKKNKQIYTTTDDYDKENIKAKHL 315
DB 230 TFRYGGYFRPGSVRLPSOTONLYNAOGNIIAKGIYDSESTTYYTFTFNVDQYTNISGSE 289
OY 316 KILSYIDKSKVPPNNNTKLDVEYKATLSSVKNKTITVEYQRENERTNLQSMFNTIDTKNH 375
DB 290 EOYAFKREKRENTATDKATYPRVEVTLGNDKYSKNYIVD---GNOKGOQLSSTNYINNEDL 346
OY 376 TVEQTYIYN-PLRYSAKETNV-NISGNGDEGSIIDDSIIKVKYKVDNON-----LPD 427
DB 347 SRMNTYVYVNOPKRTYKKEFTVNTLT-----GYKFNDAKFKIYEV-TNQNQVDSFTPD 400

```

Db 180 QGGTNAVDRKVFHESNIDIAIDKGVHNOTTGTETEMATSSDVLKLIKANTTTIDSVKESGDTFT 239

0y 35 GLEKSEERTSTINVDENATFLQKPDDHTLHEEVSSVES-----SNSIDTA 90  
 : : : : : : : : : : : : : : : :  
Db 45 CHEAKAETHGELNQSNETT-----APSEN-----KTKKVDSRLKDNTQTATAD 92

```

OY 91 OPHSTTTINREBSVQTSNDVEDSHSDPANSKIKESNTESSKEENTIQPKVKEDSTSQ 150
Db 93 QPKVT-----MSDAITYKEHSSNMOS-----PONATANOSTTK 125
OY 151 PEGYTNIDEX---ISNODELLNPEINERYENKARPLSTTSAOSPIK-----RYTVNOLOA- 201
Db 126 TSNVTTNDKSTSTYSNFTDKSNL-----TQAKDVSTTPKTTTIKPRTLRMAVNTVAAP 179
OY 202 EOGSNVHLIKVT--DOSITEGY-DSEGVYK--AHDAENLIDYTFEVDKYSQDTMT 256
Db 180 QPGTNVNDKWFHFSNIDAIIDKGHVNOQTGKTEEFWATSSDVLKUKANYTTIDSVKEGDTFT 239
OY 257 VDDKNTVPSDLTDSFTTPIKIDKSGEIIATGYDNKNKOITATFTDTYKVENIKAHLK 316
Db 240 FKIGQYRPGSVRLPDSOTONLYNAGNIIKAGIYDSTNTTTTITFTYVQYTNVRSFE 299
OY 317 LTSYIDSKVPNNNTKLDVEKTKALSSSVNKTITVEYQRPENRNTANLQSMFTNIDTKNHT 376
Db 300 QVAAKAKRKNATTDKTAVKMETVLTGNDIYSEIIVDY---GNKKAQPLISSTNYINNEDLS 356
OY 377 VEOITTIYN-PLRISAKETNV-NISGNDEGSTIIDSTTIKVKYGVGNOL-----PDSN 429
Db 357 RNMAYVNOPKNTYTKOTFEVNTLP-----GKFPNNAKNFKRIYEVTDOQNVDSFEPDTS 411
OY 430 RIVYSEVEDTNDYDQOLGNNNNVNINFGNIDS--PYIIKVI-----SKYDPNKDDYTT 482
Db 412 KLRVTTQDFDY---YSNDKNTATVYDLMKGOTSSNKQYITIQJAVYPNNSSTDNGKIDY-- 466
OY 483 IQQVTVMQTTINETYGEFRFYSYDNTIAFTSSGQGGDLPPEKTYKIGDYWEDVDKDG 542
Db 467 -----TLDTOKTKYSW---SNSYVNWGSSSTANGD-----QKKYNLGDVWEDTKKG 511
OY 543 IQQNTNDEKPLSNVLVLTITYDGS-KSVRPDEGKQDFDG 562
Db 512 KQDA--NEKGIGYVYILKDSNGKELDRTTTIDENGKYQFEG 550

RESULT 7
US-09-815-242-5471
; Sequence 5471, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 5471

```

```

: LENGTH: 1021
: TYPE: PROT
: ORGANISM: Staphylococcus aureus
US-09-815-242-5471
      Query Match 14.9%; Score 449; DB 10; Length 1021;
      Best Local Similarity 24.7%; Pred. No. 5,2e-17;
      Matches 159; Conservative 115; Mismatches 247; Indels 124; Gaps 28;

QY 2 SDEEKNDY-----NNQSIINTDNNQIKKEETNNYDGIKRESDERTESTYNDENEA 55
D 37 ADASENSYTGQSDASANEKSSDSSSVNAPKTDNTNVSD-----TKTSTNNNGET 87
QY 56 TFLQKTPDDNHLTEEVKESSVSSNSIDTAQP-----SHTTINEESQTSQDNVD 111
D 88 SVAQPAQOET-----TQSTASTNATTEETPTATGATTTATQANPATQSSN 135
QY 112 SHVSDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQSGYTNDEKISQDELLN-L 170
D 136 TNAELVN---QTSNETASNDINY-----SSVNSPQNSTAKENYSTQDSTERT 183
QY 171 PINEENKARPLSTSAOPSIKRVTVN-----QLAA-----EQGSNVNHLIKYVD 215
D 184 PSN---NSAPQSTDAKSKDYNQAVNTPSAPRMAFSLAAVAADAPAGDTITN--QLTD 238
QY 216 QSTIEGYDSDSGVKAHDAENLIDYVTEBEDVKVSGDTMYVDIDKNYPSDLTDSFTIP 275
D 239 VKVT---IDSGTTVYPHQAGYKLVNGFSPNSAVKGDFFKLTVPFELNLNGVSTAKVP 295
QY 276 KIKDMSGELIANGTDNNKNOITTYFTDYVDKYEENIKAKLITSYIDKSKVNNNTKL-D 334
D 296 PIMVGD-OVLANGVIDS-DGNYITFTDYVDKRENTANTPAIID-----PENVTKGN 349
QY 335 VEYKTAAL--SSVNTKITTEYQRPNNRRTANLOSMTNIDTKHNTVEQTIYINPLRYSAKE 392
D 350 VLTTCIGTNTASKVVLIDYEKQGFHNLSIKGTIDQIDKNTNTYQTIYVN-----SG 404
QY 393 TNV-----NISGNDESGSTIIDSTIIKYKVGDNQNLPSNRIRYDSEEDVYNDY 445
D 405 DNVTVPALTLGNLIPMTKSNALIDAKNTDILKYRV-DNANDLSSEYVNPSPDEEDYNOVR 463
QY 446 AOLGNNDVNIWF---GNIDSPYIIKYSKDPKNKDQTTIQOTFTMQTITNEYGEF- 500
D 464 ISEPNANOYKVEFPDDDOITPTIYVNVGHIDPASTG-----DIALKSTFYGDSNFI 517
QY 501 -RTASVDNTIAESTSGSQGQ--DLP--PEKYTKIGDY--VMEVDYKD----- 541
D 518 WSMMSWMDNVAANNGSGSDGDIDKPYVPQDPDEGEIIEPIEDSDSDPDSGSDSDSNDSD 577
QY 542 ---GIQNTNDNEKPLSNVLVTLTYVPDGTSKSVYR--DEDGKYOED 581
D 578 GSDSGSDSTSDSDSDSASDSDSASDSDSASDSTSDSDSASDSD 622

RESULT 8
US-09-815-242-12544
: Sequence 12544, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078

```



```

QY 485 QVTMOTINETYGERFA--SYDNITAFSTSSGOGG-----DLPEKTYKIG 531
DB 440 --LEFQTHLEGYNYTYTLTKWNGVAFYSNNAOGDKLKEPIIETHSTPIELEFKSE 497
QY 532 DYWEDVDKDGIQNTDNNEKPLSNVLTLT-----YPDGTSKSVRTDED 575
DB 498 PVEKHELTGTIESNDS--KPIDEFYHTAVEGAEHAGET---IETDED 542

RESULT 10
US-09-815-1242-13140
Sequence 83140, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13140
LENGTH: 767
TYPE: PR1
ORGANISM: Staphylococcus aureus
US-09-815-242-13140

Query Match 12.5%; Score 376; DB 10; Length 767;
Best Local Similarity 23.4%; Pred. No. 3.2e-13;
Matches 18; Conservative 110; Mismatches 217; Indels 124; Gaps 22;
QY 35 GIKRSRDRSTETTNVDENATFLQKTPQNTHTLEEVESSESSNSIDTAQ---- 90
DB 30 GOEKAASDQNTTVEES-----GSSATESKASHTOTTNNVNTIDETQSYSA 78
QY 91 ----OPSHHT-INREESVOT--SDNVEDSHVDPANSKIKESNTESGKEENTIQPKVK 143
DB 79 TSTEDPSQSTOVTEEPKPKVOAPKVTSTRY-DLPSEKVALKETGTQVD--IAQPSNV 135
QY 144 EDSTTSPSGYTNIDKISNODELLNLPINEYENKARPLSTSAOPSIKRYVNOQLAEO 203
DB 136 E-----IKPRKRSIDVTAVAEK 153
QY 204 -----GSNNHLIKVTDOSTTEGYSYKRAHDAENLIYDVFEDDKVKSGDT 254
DB 154 EYVEETKATGTDVTKKYEVEEGSEIYGHKODTNVNPNNHMERVTLKRYKMGEGIKAGDY 213
QY 255 MYVDIDKNTVPDLTDSFTIPKIKDNGSEIATGTYNKNKQULYFTTDDYDKYENIKAH 314
DB 214 FDTTLSDNVETHGISTLRKPKPEIKSTDGQVMAETGIIIGERK-VYITKEIVOEKKDLTAE 272

```

```

QY 315 LKLTSTYIDSKVPPNNNTKLDVEKFTALSSVNTKITVEY---ORPNERNRANLQSMFTNID 371
DB 273 LSLNLTIDPTTYQKQNO--NVEYKLGFTYVSKFTENIQYLGAVDNNQVTAN-----GRID 326
QY 372 TKNHT---VEOTIYINPLRYSAKETNVNISGNDSESTIIDSTIIRKVKVGNONLPDS 428
DB 327 TLNKVGGKSFHFAYMKPNNQSL--SSVTVYQGTGKNGKPPVNNPTVKYVYKHISDDLAE 384
QY 429 --NRIYDSEYEDVTDVYAQLGNNDVNINFGNIDSP--YIIRKVISKIDPNKDYTTIQ 484
DB 385 VYAKLDVSKFEDVFDNMSLDFPDNGYSLSLNFNNLDQSKNVIKYGYSNASN----- 439
QY 485 QVTMOTINETYGERFA--SYDNITAFSTSSGOGG-----DLPEKTYKIG 531
DB 440 --LEFQTHLEGYNYTYTLTKWNGVAFYSNNAOGDKLKEPIIETHSTPIELEFKSE 497
QY 532 DYWEDVDKDGIQNTDNNEKPLSNVLTLT-----YPDGTSKSVRTDED 575
DB 498 PVEKHELTGTIESNDS--KPIDEFYHTAVEGAEHAGET---IETDED 542

RESULT 11
US-09-813-820-7
Sequence 7, Application US/09813820
Patent No. US20020102262A1
GENERAL INFORMATION:
APPLICANT: Hook, Magnus
Patel, Joseph M.
House-Pompeo, Karen
Shanani, Narayana
Symersky, Jindrich
TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/813,820
FILING DATE: 22-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/856,253
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: TAMK.193
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 345 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-820-7

Query Match 11.6%; Score 349; DB 10; Length 345;
Best Local Similarity 30.8%; Pred. No. 3.6e-12;
Matches 102; Conservative 61; Mismatches 136; Indels 32; Gaps 15;

```

Qy 213 VTDQ--SITEGYDDSEGVAKAHDAENLIVTFEVDKVKSGDITMTVIDKNTVPSDLT 270  
 Db 25 ITNOLNLTAVGI-DSGTVPYPHQAGVYKLNJGFSVPNSAVKGDTEKTIPTPKELNLNGVTS 83  
 Qy 271 SPTIKRONGSEIATGTYDNKNKQITFTFDYDKVYNIAHLKLTYSIDKSKVPPNN 330  
 Db 84 TAKVPPIMAGD-QVLANGYIDS-DGNVITFTDYVNTKDVATATLMPYID-PEV 137  
 Qy 331 TKL-DVEKTLALS--VNKTIVYORPENNTANLQSMFTYEDTKNHTVEQTIYINPL- 386  
 Db 138 KTGAVTLATGISTTANKTVLVDEKYEKFEYNLSIKGIDGDKTNNYRQTIYVPSG 197  
 Qy 387-RISKETNNISGNGDESGTIIDSTIIKVKVGDNONLPSNRIDYSEVEDYNDY 445  
 Db 198 DNVAIPVLGNLKPNTDSNALLIDQNTSIIKVKYKVDNADLSESEYV-NEPNEFDVTSVN 256  
 Qy 446 AQLGNNNDVNIENG---NIDSPYIIKYSKYPN-KODYTIIQOTYMTQITINEYTG- 499  
 Db 257 ITFPNPOYKVEFTNPPDOITFTYIVVNGHIDPNKSGD-----LALRSTLYGNSNI 309  
 Qy 500 -FRITASYDNTIAFTSSGQOG-DLP--PEK 526  
 Db 310 IWRSMWMEVAFNNGSGSGDIDKPVPEQ 340

# RESULT 12 US-09-815-242-5456

; Sequence 5456, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes In  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5456  
 ; LENGTH: 978  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus aureus  
 ; US-09-815-242-5456

Query Match 10.6%; Score 319.5; DB 10; Length 978;  
 Best Local Similarity 23.2%; Pred. No. 4.9e-10;

Matches 119; Conservative 95; Mismatches 239; Indels 61; Gaps 19;

Qy 35 GIEKRSERTESTVNDENATFLQKTPDN---THLFEVKESSSVSSSSIDTAQ 90  
 Db 7 GQDKRAAASEKTTVEEN-----GNSATDNKTSQTATATVNNIEEQSYNAI--ATE 59

Qy 91 QPSHTT-INREESVOTSDNVEDSHVSDPANSKIKESNTEGKEENTIEQPN-KYNEGST 148  
 Db 60 QPSNATQVTEBAPRAVO-----APQAAQAPANLETVAEEVYKAEAKPOVETQ 109  
 Qy 149 SPPSG---YTINDEKISNDELINLPINYEKAKPLSTSKPSTKRT-----VNO 198  
 Db 110 QNNSDQKQVLDLTPKRAQONQAAETQV---EVAQPRVYSESNPRTBSADVEAKEASD 165  
 Qy 199 LAEQQSNVNLHILKLYTDSITGEGYDDSEGVAKAHDAENLIVTFEVDKVKSGDTMFD 258  
 Db 166 AKVEGTIDVTSKVYTESGSTEAPQGNK---VEPHAGQRYVLKYLKFEGLKRGDIFDT 222  
 Qy 259 IDKNTVPSDLTDSFTPIPKIKDNGSEIATGYDNKNKQITFTFDYDKVYNIAHLKLT 318  
 Db 223 LSNNTVNTGVSTARKVPEIK-NGSVYMATGQLG-NGKIRYFTFDYDKVYATDLEIN 280  
 Qy 319 SYIDSKVPNNNTKLDVEKTLSSVN-----KITIVYORPENNTANLQSMFTNIDPK 373  
 Db 281 LFIDERTVQSNQ-----QOITITSLNDKETNTLPIEYNPQVSNYSANVNGSIETFDKG 334  
 Qy 374 NHTVEQTIYINPLRYSAKETNVNISGNGDESGTIIDSTIIKVKV-GDNOLPDS--NR 430  
 Db 335 NNRTFHVAYIKP-QNGHSDSVSTIGTLQSKADGAKATVYEVYKADANLPOSVYAN 393  
 Qy 431 IYDYSEEDYVND--DYAQLGNNNDVNIENGNDSPYIIKYSKYPNKKD--DYTTIQOT 486  
 Db 394 VSDSSMFDVFEEMKDKLVENNGVYKLDIEKLESYVIHYGEYLSGSDQVNPRTHMFG 453  
 Qy 487 VTMTTINEYTGFEFTASDNTIAFTSSGQOG 520  
 Db 454 YPEQOKYYTTHLGYKLTWNGLVLSNKAQDG 487

# RESULT 13

US-09-815-242-12686  
 ; Sequence 12686, Application US/09815242  
 ; Patent No. US20020061569A1

; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes In  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12686  
 ; LENGTH: 1001  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus aureus  
 ; US-09-815-242-12686

[illegible]

```

NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5797
LENGTH: 1018
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5797

Query Match      10.2%, Score 308.5; DB 10; Length 1018;
Best Local Similarity 21.3%, Pred. No. 2e-09;
Matches 124; Conservative 118; Mismatches 242; Indels 97; Gaps 25.

QY 35 GIEKSEDRTESTTVDNDEAATFLCKTPQDN-----THLTFEEKSESSVSSNSSIDPAQ 90
   | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 30 GODKFAAASEQTTTVEEN-----GNSATDNKTSETOTTATVNHIIETOSTNATV--TE 82
QY 91 OPSHTT-INREE---SVOTSDNVEDSHSDPANSKIKESNTFSGKEENTIEDPNKVEDS 146
   ||| : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 83 OPSNATQVTTTEAPRAKVAQAPQAPANIEETVAEEVKE-----EAKPOVKEET 130
QY 147 TTSGPSG---YTNIDEKISMODELNLPINEYENKARPLSTSAOPSIKRV-----V 196
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 131 OSODNSGQORQVDLPKPKATQNOVAETQV-----EVAQPTASESKPRVRSADVAEAKEA 186
QY 197 NOLAEQGSNNVHLLKYVDQSTIEGVDSEGIKHAIDANILYDTVFVEDDKYKSGDGMT 256
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 187 SNAKVEGTGDVYSKYTVEIGSI-EGHNNTNKY-EPHAGGRAVLKYYLKFENGHLQGDYFD 244
QY 257 VIDKNTVPSDLTDSFTPIKIDNS-----GELIATGTVDNKKQITTFYFDVVKYENI 311
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 245 FTLNNVNTHGVSYTKARVPEIKNGSVMATGEVLECG-----KRYFTNDIEDKVDV 297
QY 312 KAHLLKTSYIDSKYPNNNTKLDVEKXTALSSVN-----KTTVEYQRPNEFTANLQSM 366
   ||| : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 298 TAELEINLEFDKRYOTGN-----QITSTLENEOTSKELDKYKKGIDIGNYANLNGS 351
QY 367 FFINIDTKNHTVQOTIYINLKRYSAKETNNNISGNGEGSTIIDSTIIIVYK-VGNOML 425
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 352 IETFKANNRFSHVAFIKP--NNGKTTSVTVGTGLMKGSNQNQNOCKVRIRFEYLGNNEDI 409
QY 426 PDSNRY-----DYSEVEDTNDYDQLG--NNNDVNIINGNIDSPYIIKIVSKYDPNKD- 478
   | : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 410 AAS--VYATFTTDSKKEVTSNMSGULNQNNTSYSLNIENTDKTYVNHVHDGYLNGTDE 467
QY 479 -DYTTIQQTVTMOTITNEYTGEFRASDYNTIAFSTSSGOGGDLPP-----EKTYKIGD 532
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 468 VDFRT-QVGHPRDQLKYYUDRGYTLTMDNGLVLNKNKANGKNRPIIONNKFYK--- 523
QY 533 YWVEDVDKDGIONTNDNEKPLSNVLTLLTPDSTSKSVKRD 573
   || : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 524 ---EDTIKETLNGQYDKN-----LVTTFEEEDSSSTLID 555

RESULT 15
US-09-815-242-12838
/ Sequence 12838, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes In
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/

```



```

1 PRIOR APPLICATION NUMBER: 60/206,848
2 PRIOR FILING DATE: 2000-05-23
3 PRIOR APPLICATION NUMBER: 60/207,727
4 PRIOR FILING DATE: 2000-05-26
5 PRIOR APPLICATION NUMBER: 60/242,578
6 PRIOR FILING DATE: 2000-10-23
7 PRIOR APPLICATION NUMBER: 60/253,625
8 PRIOR FILING DATE: 2000-11-27
9 PRIOR APPLICATION NUMBER: 60/257,931
10 PRIOR FILING DATE: 2000-12-22
11 PRIOR APPLICATION NUMBER: 60/269,308
12 PRIOR FILING DATE: 2001-02-16
13 NUMBER OF SEQ ID NOS: 14110
14 SOFTWARE: FastSeq for Windows, Version 4.0.
15 SEQ ID NO 12838
16 LENGTH: 1018
17 TYPE: prt
18 ORGANISM: Staphylococcus aureus
19 09-815-242-12838

```

Query Match	10.2%	Score 308.5	DB 10	Length 1018
Best Local Similarity	21.3%	Pred. No. 2e-09		
Matches 124	Conservative 118	Mismatches 242	Indels 97	Gaps 25

```

OY 35 GIERKSEBERTESTINVENENETFLQKPIPODI---TUTTEEYESSSSSSSIDTAQ 90
D 30 GODDEAASAASEKTTIVVEEN-----GNSATDNKSTETOTTTATNHIIEFOSYKATV--TE 82
OY 91 QPSHTT-INREE---SVQTSDNVEDSHVSPFANSKIKESNTSGREBENTIEOPNKYKEDS 146
D 83 OPSNATQVTTTEAPKRAVOAPOTAQAPANIEFVEKVEA-----EAKQVYKETT 130
OY 147 TTTSQPSG---YTINDEKISNDELINLPINIEYKAKAPLSTTSQPSIKRYT-----V 186
D 131 QSDNSNGQROVDLTPPKATONOVAETQV---EVAOPRTASEKPRVTSADVAEAKEA 186
OY 197 NQLAAEGSNVNHILIKYTDOSITEGYDQSEGVKAHAENLIYDVTPEYDVKYSGDTMT 256
D 187 SNAKVEITDVTSKYVEIGSI-EGHNNTKVV-EPHAGQRAVLKYKKAFENGHOGDYFD 244
OY 257 VDIKNTVPSDLTDSFTPIKIDNS---GELIATGTYNKKNQITVFTFDVYDKYNTI 311
D 245 FTLSNVNVTGCVSPARKVPEIKNSVYMAETGEVLEGG-----KINTVFTNDEIDKVDV 297
OY 312 KAHKLTUYIDKSKVPNNNTKLDVEYKTALSSVN-----KTIYEVORPENRNTAIQSM 366
D 298 TAELEINLFIIDPKVYOTNGN-----QITISLNEEOTSKELDYKYXDQIGNYAANLNGS 351
OY 367 FTNIDTKHTEVEQTUYIINPLRYSAKETVNVNSGNGDGSTIIDSITIIKYK--GVDGNQL 423
D 352 IETFEKANNRSHVAFLIP--NNGKTTISVYTGILMGKSNONGNQPKVRIFEIYGNEDI 409
OY 426 PDSNRUY---DYSEYEDVTDNDYDAOLG--NNNDVNINFGNIDEPYIIKYISKYDPMKD- 478
D 410 AKS--VYANTJTDTSKFEKVTSSMGSGLNLQNGSYSLNIEHLDKTYVYHHDGVELNLTDE 467
OY 479 -DYTTIQOVTVMQTTINETYEEFRTASDNTIAFSTSSGQOGFLP-----EKTYIGD 533
D 468 VDFRT-QWVGHPEQLYKYYRQGYTLTMDNGLVYYSKNAMNGKNCPIIONNKFEYK-- 523
OY 533 YVMEVDKDGLONTNNDNKEPLSNVLVYTLTYPDGTSKSRPD 573
D 524 ---EDTIKETILGQYDKN-----LVYTFVEEYDSSLTLD 555

```

Search completed: March 14, 2003, 13:07:57  
Job time : 12.6918 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:57:01 ; Search time 11.8208 Seconds  
(without alignments)  
4733.207 Million cell updates/sec

Title: US-09-147-405B-13

Perfect score: 3013  
Sequence: 1 SSDEEKNDVYINNOSINTFDD.....PDGTSKSVPTDEGKYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2 T30214	fibrinogen-binding
2	1144.5	38.0	1166	2 T28680	fibrinogen-binding
3	1137.5	37.8	1141	2 E89824	hypothetical prote
4	576	18.1	1315	2 T28679	fibrinogen-binding
5	559	18.6	1385	2 D89824	hypothetical prote
6	489	16.2	953	2 C89824	hypothetical prote
7	452	15.0	933	2 S41539	fibrinogen-binding
8	452	15.0	989	2 D89852	fibrinogen-binding
9	397.5	13.2	877	2 F90070	Clumping factor B
10	376	12.5	940	2 S19702	fibrinogen-binding
11	344.5	11.4	961	2 G90053	hypothetical prote
12	326	10.8	1038	2 H90053	hypothetical prote
13	300.5	10.0	1018	2 A32192	fibrinogen-binding
14	239	7.9	1039	2 T30856	protein F2 - Strept
15	225	7.5	1463	2 T30290	AAS surface protei
16	216.5	7.2	1959	2 AG1085	hypothetical prote
17	216.5	7.2	3394	2 T18501	hypothetical prote
18	212.5	7.1	1125	2 E90598	membrane nucleas
19	211.5	7.0	2206	2 G71611	hypothetical prote
20	208.5	6.9	1072	2 A86827	hypothetical prote
21	208	6.9	1127	2 T28317	ORF MSV156 hypote
22	208	6.9	3724	2 T18427	hypothetical prote
23	207.5	6.9	1711	2 T18429	hypothetical prote
24	207.5	6.9	2401	2 T28676	rhodopy protein -
25	207	6.9	769	2 F89870	serine proteinase
26	207	6.9	1939	2 T18372	repeat organellar
27	207	6.9	4550	2 T18440	hypothetical prote
28	206.5	6.9	821	2 S67087	hypothetical prote
29	206	6.8	1302	1 JG6009	surface-located me

30	205.5	6.8	1365	2 T30822	
31	205.5	6.8	2510	2 T28160	
32	205	6.8	2481	2 D90011	
33	204	6.8	3844	2 T18402	
34	203	6.7	537	2 A23770	
35	202.5	6.7	1997	2 F71607	
36	202.5	6.7	2269	2 T28677	
37	202	6.7	1093	2 T18275	
38	201.5	6.7	4688	2 F82885	
39	199	6.6	665	2 B71609	
40	198.5	6.6	807	2 B71605	
41	198.5	6.6	1805	2 T02712	
42	198.5	6.6	3147	2 T18674	
43	198	6.6	1308	2 E71622	
44	197	6.5	940	2 A01374	
45	197	6.5	2523	2 T18477	

## ALIGNMENTS

## RESULT 1

T30214

fibrinogen-binding protein - Staphylococcus epidermidis

C:Species: Staphylococcus epidermidis

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000

C:Accession: T30214

R.Nilsson, M., Frykberg, L., Flock, J.I., Pel, L., Lindberg, M., Guse, B.

Infect. Immun. 66, 2666-2673, 1998

A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.

A:Reference number: Z20781; MUID:98261511; PMID:9596732

A:Accession: T30214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1092 <NLI>

A:Cross-References: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAN76638.1

Query Match 100.0%; Score 3013; DB 2; Length 1092;

Best Local Similarity 100.0%; Pred. No. 6.6e-123;

Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SSDEEKNDVYINNOSINTFDDNQIIKKEETNNYDGIKESSEDETESTTVDENEATFLOK	60
DB	75	SSDEEKNDVYINNOSINTFDDNQIIKKEETNNYDGIKESSEDETESTTVDENEATFLOK	134
QY	61	TPQDNTHLTEEEYKESSEYSSNSIDTFAQPSHTTINREESYQTSQDNYEDSHVDFANS	120
DB	135	TPQDNTHLTEEEYKESSEYSSNSIDTFAQPSHTTINREESYQTSQDNYEDSHVDFANS	194
QY	121	KIKESNTESGKEENTIEQPKYKEDSTTSQPSGYTINDEKISQDELNLPINEYENKAR	180
DB	195	KIKESNTESGKEENTIEQPKYKEDSTTSQPSGYTINDEKISQDELNLPINEYENKAR	254
QY	181	PLSTSAQPSIKRYTVNOQLAEGSVNHLIKTQDSITFGYDSSGCVYKAHAENLIYD	240
DB	255	PLSTSAQPSIKRYTVNOQLAEGSVNHLIKTQDSITFGYDSSGCVYKAHAENLIYD	314
QY	241	VFEVDDKYSQDPTMTVDIDKNTVPSDLGFSFPIKIDKSGEIIITGTVDNKNKOITYT	300
DB	315	VFEVDDKYSQDPTMTVDIDKNTVPSDLGFSFPIKIDKSGEIIITGTVDNKNKOITYT	374
QY	301	FTDYVDKYENIKAKLITSYIDSKVPNNNTKLDVEYKATLSSVNTTIVEYQRPENRT	360
DB	375	FTDYVDKYENIKAKLITSYIDSKVPNNNTKLDVEYKATLSSVNTTIVEYQRPENRT	434
QY	361	ANQSMFTNIDTNNHYEQITINPLRYSAKKEYNVNSGNGDGSFTIDSTIIKKYKVG	420
DB	435	ANQSMFTNIDTNNHYEQITINPLRYSAKKEYNVNSGNGDGSFTIDSTIIKKYKVG	494
QY	421	DNQNLDPDSNRITYSEYEDVTNDYQALQGNNDVNIFFGNDSPYIIKIVISKYDPKNDY	480
DB	495	DNQNLDPDSNRITYSEYEDVTNDYQALQGNNDVNIFFGNDSPYIIKIVISKYDPKNDY	554



## RESULT 4

fibrinogen-binding protein homolog - *Staphylococcus aureus*

T28679

C:Species: *Staphylococcus aureus*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000

C:Accession: T28679

R:Johansson, E.; McCrear, K.; N. Eldholm, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.

Microbiology 144, 3387-3395, 1998

A:Title: Three new members of the serine-aseptate repeat protein multigene family of *S.*

A:Reference number: 220510; MUID:99098700; PMID:9884231

A:Accession: T28679

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1315 &lt;OS&gt;

A:Cross-references: EMBL:AJ005646; NID:e1318791; PID:e1318792; PIDN:CAA06651.1

C:Genetics:

A:Gene: sdrc

## Query Match

Best Local Similarity 19.1%; Score 576; DB 2; Length 1315;  
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 1 SSDEKNDV---INNOSINTDDNNQIKKEETNNYDGIKRSERSTSTNVDE-NEAT 56  
 DB 55 STNKLNEATTASASDNQSDKVDQMLQNEQNTKNDKOKEMVSSQGNFTTSNGNKLKE 114  
 OY 57 FLQKTPDNTHTLEEEVSESSVESNSIDTAQPSHTTINRESVQTSVDNVEDSHYSD 116  
 DB 115 SVQSTGKRVETSTAKSDEQSPKSTNEDLNTKQ---TISNQAQ-PDLQEKSVYN 168  
 OY 117 FANSKIKESNTESGKEENTIEOPN-KVEDSTTSQPSGYNTI-DEKISMODELL-NLEPIN 173  
 DB 169 ---VQPTNEBKVKVD-----AKTEST-----LWXSDAIKSDELVLVNNNS 209  
 OY 174 EYENKARPLSTSAOP---SIKRYTVQALAEQGSNNHLIKYDQSTIBETDSEGYI 229  
 DB 210 NNNENNADILPKSTAPKRLNTRMRLAAVQPSSTEAKNNDLITSTVTLVVDADKNNKI 269  
 OY 230 KAHAEENIYDVFEDVDKYSGGDTMTVDIDKNV-----PSDLDSFTPIKID-NSG 282  
 DB 270 PRQDLSLKSQIT--VDKRVSGDFTIKY--SDIVQVGLNPEDIKN---ISDIDPNNG 323  
 OY 283 EIAATGYDNKKKQITFTFTVDYDKYENIKAHKLTSYIDKSKVPNNTKLDEYKTAIS 342  
 DB 324 ETIAAKHDFANMLITFTYDVRFSVQMGINSYIMADATIP--VSKNVEENVTIG 381  
 OY 343 SVNKTIYEVORP--ENRTPANLQSMFTNIDTKNHT-----VEQTIYINPLKSAKE 392  
 DB 382 NTTTETANIQYPDYVVEKNSIGSAFT--ETVSHVGNKENGYYKQTIYVNPSENSLTN 439  
 Y 393 TNVNI-----SGNGDEGSIIDSTLIKYKVGDNQNLDPDSNRIYYS--EYEDVINDDA 446  
 DB 440 AKLQAQVHSSYPNINIGQINNKVDITIKIYQPKGYTL--NGYDVNKLDELVDN-QYL 495  
 OY 447 Q---LGNNDVNIINGNDSPYIIRVISKYDPNKKDYTIQOTVMTQITINEYGEFFTA 503  
 DB 496 QKITGDNNSAVIDFGNMDASVYVWNTKFOYTNSESPTLVOMATLSTSGN-----KSV 549  
 OY 504 STDNTIAFSTSSGGQGLPEPKTYKIGDYVEDYDKGQIQNTNDEKLSNVLTLYLP 563  
 DB 550 STGNALGFTNNSGGAG---QEVYKIGYVWEDTNKNGVQL--GEKGVGVVTVV-FD 602  
 OY 564 DGTSKSVR---TDEDGY 578  
 DB 603 NNTNTRKGEAVTKEDGSY 620

## RESULT 5

hypothetical protein sdrc [imported] - *Staphylococcus aureus* (strain N315)

DB9824

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: DB9824

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;

ma, A.; Mizutani, U.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu,

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: DB9824

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1385 &lt;KUR&gt;

A:Cross-references: GB:BA000018; PID:g13700454; PIDN:BA041751.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: sdrc

## Query Match

Best Local Similarity 18.6%; Score 559; DB 2; Length 1385;  
Matches 177; Conservative 117; Mismatches 235; Indels 86; Gaps 27;

QY 1 SSDEKNDV---INNOSINTDDNNQIKKEETNNYDGIKRSERSTSTNVDE-NEAT 56  
 DB 55 STNKLNEATTASASDNQSDKVDQMLQNEQNTKNDKOKEMVSSQGNFTTSNGNKLKE 114  
 OY 57 FLQKTPDNTHTLEEEVSESSVESNSIDTAQPSHTTINRESVQTSVDNVEDSHYSD 116  
 DB 115 SVQSTGKRVETSTAKSDEQSPKSTNEDLNTKQ---TISNQAQ-PDLQEKSVYN 168  
 OY 117 FANSKIKESNTESGKEENTIEOPN-KVEDSTTSQPSGYNTI-DEKISMODELL-NLEPIN 173  
 DB 169 ---VQPTNEBKVKVD-----AKTEST-----LWXSDAIKSDELVLVNNNS 209  
 OY 176 ENKARPLSTSAOPSIKRYTVQALAEQGS---NVNHLIKYDQSTIBETDSEGYIKA 231  
 DB 212 ENNADILPKSTAPKRLNTRMRLAAVQPSSTEAKNNDLITSTVTLVVDADKNNKI 271  
 OY 222 HDENLIIYDVFEDVDKYSGGDTMTVDIDKNV-----PSDLDSFTPIKID-NSGEI 284  
 DB 272 QDYLKSQIT--VDKRVSGDFTIKY--SDIVQVGLNPEDIKN---IGDIDPNNGET 325  
 OY 285 IATGYDNKKKQITFTFTVDYDKYENIKAHKLTSYIDKSKVPNNTKLDEYKTAISV 344  
 DB 326 IATAKHDFANMLITFTYDVRFSVQMGINSYIMADATIPV--KQVFPSTVIGQ 383  
 OY 345 NKTIYEVORP--ENRTPANLQSMFTNIDTKNHT-----VEQTIYINPLKSAKETN 394  
 DB 384 ITTTADITTPAKYKADNNSIGSAFT--ETVSHVGNVEDPGYVNVYVNPMDKLGAK 441  
 OY 395 VNISNGDEGSI---IIDSSTLIKYKVGDNQNLDPDSNRIYD--SEYEDVIND--DYA 446  
 DB 442 LKVEAVHPKYPPTNIGQINQNTNIRIKIYRVEGYTL---NGYDVNKLDELVDTEDEKKN 498  
 OY 447 QGNNNDVNIINGNDSPYIIRVISKYDPNKKDYTIQOTVMTQITINEYGEFFTA 506  
 DB 499 TVGNSQVNLDFGLDSATVYVWNTKFOYTNSESPTLVOMATLSTSGN-----KSVIG 552  
 OY 507 NTIAFSTSSGGQGLPEPKTYKIGDYVEDYDKGQIQNTNDEKLSNVLTLYLPDGT 566  
 DB 553 NALGFTNNSGGAG---QEVYKIGYVWEDTNKNGVQL--GEKGVGVVTVV-FDNN 605  
 OY 567 SKSVR---TDEDGY 578  
 DB 606 NTRKGEAVTKEDGSY 620

## RESULT 6

hypothetical protein sdrc [imported] - *Staphylococcus aureus* (strain N315)

C9824

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: C9824

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;  
 ma, A.; Mizutani, U.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu,  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.



```

Db 51 SNEKSNDSSSVSAAPKPTDDTNV-----SPTKTSNTN---NGETSVQA 91
Qy 61 TPQDNTHLTEEVKSSSVSSSSSIDTAQOPSHHTTINEESVQSDNVEDSVSPFAS 120
Db 92 NP-----AQOETTSSTNATTEETPVNGEATTTTNOAMPATQOS--SNTNAELV- 143
Qy 121 KKESTESGKEENTIEQNKVKEDSTQSPGYTIDEXISNOELLN-LPINEYENKA 179
Db 144 --QTSNETTSNDTNTV-----SSVNSPONTNAENVSTQDTSTATPSN---NES 189
Qy 180 RELSTTSA-----OPSIKRVYNQOLA-----EOGSVNHILIKYTDOSITEGYD 224
Db 190 APOUNTDAKSDVYSAOVNSTPRMRFRSLAAVAADAPAGTDTTN--QLTDVAVT---ID 244
Qy 225 SEGVAKAHDAENLIYDVTEFVDKYSQDPTVTDIKNTVPSDLTDSFTIPKIKDSGCI 284
Db 245 SGTIVYPOAGYVKLNYGFSVPNSAVKGDPTKTVPEKELNLNGVSTAKVPPIMAGD-QY 303
Qy 285 IATGTYDNKKNQITTYFTFYVDKYEKIKAKLSTYIDSKVYNNNTKL-DVEYKTAAL-- 341
Db 304 LANGVIDS--DGNVYTFYFDYVNDKENVNTANTIMPAYID---PENVTKGWVLTGIGT 358
Qy 342 SSVNTKITVEYORPENRFRANLOSMTNIDTKNHTVEQTIYINPLRYSKAEYV----- 395
Db 359 NTASKTVLIDYEKYGOFNLSTIKGTIDQIDKTNNTYRQTIYVNP-----SGDKVVLPA 413
Qy 396 -NISGNGDEGSIITIDSTIIKVKYVGNQNLPSNRIYDSEYEDVTNDYDQAOLGNNDY 454
Db 414 GNLIPTKSNALIDAKNTDIKYRV--DNANDLSESYVNSDPEDEYTNQVIRISFPANQY 472
Qy 455 NINF-----GNISPTIIRIKISKYDPKDKDYTTIQOVTVMQTTINEYGER--STASYDN 508
Db 473 KVEFPTDDQITTPYIVVNGHIDPASTG-----DLAKSTYIGDSNFTIMRSMWDNE 526
Qy 509 IAFSTSGOGQOG-DLP--PEKTYKIDY--VMEDVVDK-GIQMTNDN 549
Db 527 VAFNNSGSGDGDIGKPVPEQPPPEGEIIEIPDSDSPOSDSGSDS 573

RESULT 9
Clumping factor B [imported] - Staphylococcus aureus (strain N315)
F90070
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90070
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C.; Shiba, T.; Hatori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; M0ID:21311952; PMID:11418146
A:Accession: F90070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <R>
A:Cross-references: GB:BA000018; PID:g13702588; PIDN:BA843728.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: clfB

Query Match 13.28; Score 397.5; DB 2; Length 877;
Best Local Similarity 24.5%; Pred. No. 4.3e-10;
Matches 143; Conservative 81; Mismatches 207; Indels 153; Gaps 23;

```

```

Db 143 NSQVNDKKT-----NDANSIATNSELKNSQT-----LDLP-----Q 173
Qy 179 APRLSTTSAQ-----PSIKRVYNQOLAEO-----GSNNHILIKYTDOSTIEGYD 225
Db 174 SSPORTISNAQGSKPSVFRRAVRSLSVAEPVYNAADAKGTNN-----DKYASNFKE 227
Qy 226 EGVYKHAHDAENLIYDVTEFVDKYSQDPTV-----DID-----KNTVPSDLTDS 271
Db 228 KTFEPDNGSGTFMAANFTVTDKVSQDYFTAKLPDLSLNGNDVDSNSNNTNP----- 281
Qy 272 FTIPKIKDSGEIATGYDNKKQITTYFTFYVDKYEKIKAKLSTYIDSKVYNNNT 331
Db 282 --IADIKSNNGDVYAKATDILKTYTFVFTOVNNKENINGFSJLPFTDRAKAKSG 339
Qy 332 KLDVEYKTAALSSVNTKITVEYOR-----NENRTANLOSMTNIDTKN--HYEQTIYIN 385
Db 340 -YDANINIDEMFNKRTYNSSPISGIADIKPNCANISIOIGVDTASGONTYKQTVFVNP 398
Qy 386 LRSKAEYVNVISGNGDEGSIITDS-----TIKVKYVGNQNLPS----- 428
Db 399 KORVLGNTWYIKGYDK-----IESSGKVSATDTKLRIEVDNDSKLSDSYIADPNDS 454
Qy 429 -----NRIYDSEYEDVTNDYDQAOLGNNDVNIENGNDSPYIIRIKVSKYDPNKD 479
Db 455 LKEVTDOFKNRIT-----YE-----HPVVASIKGDIKTYVVLVEGHYDNTGKN 499
Qy 480 YTT--IQOVTVMQTTINEYGERFTASVDNTIAFSTSGOGQD 521
Db 500 LKTOVIOENVDVPT-----NRDYSIFGMNENNVRYVGGSGADSD 538

RESULT 10
Fibronectin-binding protein B - Staphylococcus aureus
F90070
C:Species: Staphylococcus aureus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C:Accession: S19702
R:Joensson, K.; Stinaes, C.; Mueller, H. P.; Lindberg, M.
Eur. J. Biochem. 202, 1041-1048, 1991
A:Title: Two different genes encode fibronectin binding proteins in Staphylococcus aureus.
A:Reference number: S19702; M0ID:92111475; PMID:1837266
A:Accession: S19702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-940 <O>
A:Cross-references: EMBL:X62992; NID:g49040; PIDN:CAA44726.1; PID:g581562
C:Keywords: fibronectin binding

Query Match 12.58; Score 376; DB 2; Length 940;
Best Local Similarity 23.48%; Pred. No. 4e-09;
Matches 138; Conservative 110; Mismatches 217; Indels 124; Gaps 22;

```



Db 506 --KNGQIIQDNDPEFKEDTAKGTMSGQYDAKQIIETEENQDNTPLDIDYHFAIDEGGX- 562  
 QY 580 FDG 582  
 Db 563 VDG 565

# RESULT 13

A32192  
 fibronectin-binding protein - Staphylococcus aureus  
 C:Species: Staphylococcus aureus  
 C:Date: 31-Jul-1989 #sequence\_revision 31-Jul-1989 #text\_change 16-Feb-1997  
 C:Accession: A32192  
 R:Signaes, C.; Raucel, G.; Joensson, K.; Lindgren, P.E.; Anantharamiah, G.M.; Hoeck, M.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 699-703, 1989  
 A:Title: Nucleotide sequence of the gene for a fibronectin-binding protein from Staphylo  
 A:Reference number: A32192; MUID:89098998; PMID:2521391  
 A:Accession: A32192  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1018 <SIG>  
 A:Cross-references: GB:J04151  
 C:Keywords: fibronectin binding

Query Match 10.0%; Score 300.5; DB 2; Length 1018;  
 Best Local Similarity 21.2%; Pred. No. 8.1e-06;  
 Matches 123; Conservative 118; Mismatches 245; Indels 97; Gaps 25;

QY 35 GIEKSEDETESTVNDVEDEATFLOKTPDN-----THLFEVYKSSSVSSSSSSTIDTQ 90  
 Db 30 GQDKRAAASEQKTTVEEN-----GNSATDNKTSSETQTTATVNHIEQTQSYNAV--TE 82  
 QY 91 QPSHTT-INREE---SVQTSNDVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVEDS 146  
 Db 83 QPSNATQVTEAPKAVQAPQTAQANITVKEVYKE-----EAKPVKETT 130  
 QY 147 TTSQPSG---YTNIDEKISNDELNLPLINEXENKARPLSTSAQSIKRV-----V 196  
 Db 131 QSDNSGQDQRYDLPRKKAQVQVETQY---EVAQPTASESKPRVRSADVAEAKDA 186  
 QY 197 NOLAEQGSNNVHLKVTQSTEGVDSEGYIKAHDAENLYDYTFEVDKVGSGDTM 256  
 Db 187 SNAKVEGTGVTSKYTVEIGSI-EGHNNTNKV-EPHAGRAVLKYLKFEENGJHQGDYD 244  
 QY 257 VOIDKTVSDLTDSFTPIKIKDNS-----GELIATGYDNKKQIYTFDVKYENI 311  
 Db 245 FTLSNNVNHGVTAKKVEIKNGSYMATGEVLBEG-----KIRYFTNDIEDKVDV 297  
 QY 312 KAHLLKTSYIDSKVYNNNTKLDVEYKTLASSVN-----KITVEYQRPNEKFTANLQSM 366  
 Db 298 TALEELNLFIDPKVTQTN-----QTITSLNEQTSKELDVKKYKDGIGNYANLNGS 351  
 QY 367 FENIDKKNTEVQTIYNLAKRSAKETNNISGNGDEGSTIIDSLIKIYK-VGDNOML 425  
 Db 352 IETFNANRRFSHVAFIKP--NNGKTSVTVGTGTLKNGSNQNGQPKRVIRFEELGNEDI 409  
 QY 426 PDSNRIRY----DYSEYEDYNDYAOLG--NNNDVINENGINDISPILIKIYSYDNKPD 478  
 Db 410 AAS--VYANTTDSKKEVTSNMSGNLNIONNGSYSLNENIDKTYVHVHGDGYLNGTDE 467  
 QY 479 -DYTTIQQTVMQTTINEYGEFRTASYNQNTAFSTSSGQGGDLP-----EKYKIGD 532  
 Db 468 VQFRT-QMVGHPQLKYYUDDRGYTLTMDNGLVLNKNKANGNEKNPIIONNFEYK--- 523  
 QY 533 YVWEDYDKDGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTD 573  
 Db 524 ---EDTKEFTLGGYDKN-----LVTVVEEYDSSTLDD 555

RESULT 14  
 T30856  
 protein P2 - Streptococcus pyogenes  
 C:Species: Streptococcus pyogenes

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
 C:Accession: T30856  
 R:Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanskl, E.  
 Mol. Microbiol. 21, 373-384, 1996  
 A:Title: Protein P2, a novel fibronectin-binding protein from Streptococcus pyogenes  
 A:Reference number: 220907; MUID:97011581; PMID:8858591  
 A:Accession: T30856  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1039 <JAF>  
 A:Cross-references: EMBL:U31980; NID:g1654115; PID:g1654116; PIDN:AAC44522.1  
 C:Genetics:  
 A:Note: PRPF2

Query Match 7.9%; Score 239; DB 2; Length 1039;  
 Best Local Similarity 22.6%; Pred. No. 0.0038;  
 Matches 93; Conservative 79; Mismatches 160; Indels 72; Gaps 18;

QY 217 SITTEGIDDSSEGYKAMD---AENLYDYTFEVDKVK-----GDTMTVIDDKNTVPSDL 268  
 Db 38 SVSKRYGEQEKTSNADDFRNHAAYFKMSFELKQDKSETINPGDTEFVLQIDRLRPKGI 97  
 QY 269 TDSFTPK-IKDNSEIIATGYDNKKQIYTFDVKYENIKAHLLKTSYIDSKVYP 327  
 Db 98 SOD--TPKIIDSENSPLAIGKTYDAKTHQIYTFNTYAGLDKQVLSBELSLEKLEYL 155  
 QY 328 NNNTKLDVEYKATLSSVNTIVETVYORPNE-----NRTANLQSMETNIDTNHVE 378  
 Db 156 EENTNIDSFSTIGGEIYTKGVNVLYGNSRESNYITNGLSNVGGSIESVNTETGEV 215  
 QY 379 QTIYNPLRYSAKETNNVNSG-----NGDESTIIDDS--TIVYKVGDNQNPDSNR 430  
 Db 216 WYVYVNPNTNIPYAVNLMLGFAKRTAOGENNSLYVSQOLGYDIEVPHYRPLTSTG 275  
 QY 431 IYDYSEYEDVTNDVYAOI-----GNNDVINIFG-NID-SPYIIKIVISKYDPNKDDYTT 482  
 Db 276 V-DISRL-NLRKDLEKLQSGSTQGANRRLIDFGENLQKAFVAVYTGKADQSGKE--- 330  
 QY 483 IOQFTVMTQTINEYT--GEFRT-----ASYDNTIAFSTSSGQGG-----DLPEKT 527  
 Db 331 ----LIVGHLTSFNMWGSYKTLRPNSHVSFNEIALSPSKSGSSTSEFTKPSITVANL 386  
 QY 528 YKIGDYVEDVDKDG-----ONTDNEKPL---SNVLTLYTPDGT 567  
 Db 387 KRVQALRFKKVSTVDNPLPEAFAELRSSNGNSOKLEASSNTQGEIHFQDLTS 438

## RESULT 15

T30290  
 AAS surface protein - Staphylococcus saprophyticus  
 C:Species: Staphylococcus saprophyticus  
 C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
 C:Accession: T30290  
 R:Hell, W.; Meyer, H.G.W.; Gattermann, S.G.  
 Mol. Microbiol. 29, 871-881, 1998  
 A:Title: Cloning of aas, a gene encoding a Staphylococcus saprophyticus surface pro  
 A:Reference number: 220809; MUID:98389318; PMID:9723925  
 A:Accession: T30290  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-1463 <HEL>  
 A:Cross-references: EMBL:AJ000007; NID:e1295629; PID:e1295630; PIDN:CA03852.1  
 C:Genetics:  
 A:Gene: aas

Query Match 7.5%; Score 225; DB 2; Length 1463;  
 Best Local Similarity 20.9%; Pred. No. 0.024;  
 Matches 143; Conservative 105; Mismatches 263; Indels 172; Gaps 33;

QY 1 SPSDEKNDYINNQSI-----NTDNNQIIKKKEETNNYDGIEKRSDEPSTST--- 48  
 Db 180 NTDENSVDVANQNEPVAQNDKATETSNEDVASSDYKDDGTHSDNMSADDLTDQNESETLND 239



```

OY 49 -NVENENTFLQKIPDQDTH-----LIE--EEVNESSVSSN-----SIDTAQOPSH 94
Db 240 NAASSNEDVASSDYKQOGTHSDDNASDLDQDNEQAONKRAETSNEDVASSDYKQDTH 239
OY 95 TTIN-----REESVQTSDNVEDSHVDFANSKIKESNETSG-----KEENTIEOPN- 140
Db 300 SDANAADVAQDNESETQNDNAKNESNEDVASSDYKQDTHSDANAASDYKQDNESETQND 359
OY 141 -----KYKEDSTTS--QPSGYNNIDEKISNODELLNPLINEYENKARPLSTT 185
Db 360 KAETSNEDVASSDYKQDTHSDANASDIDQDNEQAONKRA-----ETSNEDVASS 411
OY 186 SAQPSIKRVTYNQOLAQOGSVNHLIVT---DQSTIEGYDQS--EGVTK-AHDAENLT 238
Db 412 DKQDTHSDANASDIAQDNEAATDQDQATSKEDDYVSNQDQNAKYNNITKEASTAENKY 471
OY 239 YDVTF--EVDKVKSGDTM--TVDIDKNTPSDDLTFPIPKIKDNGSELIANGTYDNKN 294
Db 472 QPATEFSKVPYKPLKRVATTSANTAVATSAVTKAETRAALPK-----S 515
OY 295 KOIYTFPDIYDKKENIKA---HLKLSYIDKSKVPNNKTKLDEYKTYTALSSVKNKITYE 351
Db 516 PKVSSINNYYRK--NNFKAPYEQDIDIANLYPLQYNYRYKQKREGIYMHDTA--NONSTITGE 572
OY 352 YQRPENENTAMLOSMFNIDTKNHTVEQTYIYNLPRYSAKETUNVNISGNDGEGSTIIDQS 411
Db 573 INYKNKNT-----SAFYHAYVDGRIIEFTANTDYLAMGA-----QPDANDRE 615
OY 412 TIIRKYKVGDNQNT.PDSNRITYDSEYEDVTNDYDQAOLGNNDVYNINFGNIDSPYIIRVIS 471
Db 616 IHVELVHTHDYDSEFARS--INNYADVA-ATNLQYGYGVLPD---SAEYDQGGYVTHQAVS 669
OY 472 KY-----DN-----KQDYATTIOQVMTQNTINEYSEFPYASTADMTIAFST- 513
Db 670 NYLGGSDHSDPHGLAANHSYDELDIOLITEKYLIKFG-----QAAMGTTSSGSTG 721
OY 514 ----SSGQGO-GDLPEEK--TYKIGDYVMEVDYDKQDQIANTNDNKEPLSNLVLTLYPDDG 566
Db 722 GTGGGTGSGNGTGPSPKSGTVK-----TENNGVGRINSK---NDGYYTTVYDQKG 770
OY 567 SKSVATDED-----GKQOF 580
Db 771 KTRDVRNQTLLVTKVSAITLQKEQY 793

```

```

search completed: March 14, 2003, 13:01:24
job time : 16.8208 secs

```

GenCore version 5.1.4-F5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 7.64875 Seconds

(without alignments)  
3155.972 Million cell updates/sec

Title: US-09-147-405b-13

Perfect score: 3013

Sequence: 1 SSDEKNDYINNQSINTDD.....PDGTSKSVRTDEGKQFQD 582

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300.5	10.0	1018	FNBA_STAU	P14738 staphylococ
2	203	6.7	537	ARP_PLAFA	P04931 plasmodium
3	202	6.7	1093	PIAK_DICDI	P54677 dictyosteli
4	196.5	6.5	1139	HMM1_MYCGE	Q09413 mycoplasma
5	192	6.4	918	YMBJ_CAEEL	P34487 caenorhabdi
6	189.5	6.3	1664	INTL_CANAL	P53705 candida alb
7	188.5	6.3	578	LIPA_MYCPU	Q50274 mycoplasma
8	188.5	6.3	1435	EBAL_PLAFC	P19214 plasmodium
9	186.5	6.2	3110	LMAD2_HUMAN	P24043 homo sapien
10	184.5	6.1	1744	TANA_XENLA	Q01550 xenopus lae
11	184	6.1	817	YG4A_YEAST	P46948 saccharomyc
12	184	6.1	1251	RBP2_PLAVB	Q00799 plasmodium
13	183.5	6.1	2009	SEC7_YEAST	P11075 saccharomyc
14	183	6.1	2452	RPB1_PLAFC	P14248 plasmodium
15	182.5	6.1	853	YCGL_YEAST	P25588 saccharomyc
16	182.5	6.1	1381	YBE7_YEAST	P34216 saccharomyc
17	182.5	6.1	1658	YME7_YEAST	Q03661 saccharomyc
18	182	6.0	1419	ALAI_CANAL	Q13368 candida alb
19	182	6.0	1460	N159_YEAST	P04577 saccharomyc
20	181.5	6.0	1093	YKDS_CAEEL	Q03563 caenorhabdi
21	179.5	6.0	1165	YNF4_YEAST	P53950 saccharomyc
22	179	5.9	1253	DSPF_HUMAN	Q09244 homo sapien
23	178	5.9	960	YMX6_YEAST	Q04279 saccharomyc
24	178	5.9	1875	YME1_YEAST	Q04455 saccharomyc
25	177	5.9	2660	YEEL_ECO57	Q08877 escherichia
26	176	5.8	1196	BXCN_CIOBO	P46081 clostridium
27	175.5	5.8	1183	CNA_STAU	Q53654 staphylococ
28	175.5	5.8	1637	MSPF_STAU	P80544 staphylococ
29	175	5.8	1147	CGAI_HELPEY	P80200 heliobacte
30	174	5.8	2022	ANT1_ONCVO	P21249 onchocerca
31	173.5	5.8	967	YB5Q_CIOAB	P33747 clostridium
32	173	5.7	1310	YB35_SCHPO	Q14340 schistosach
33	172	5.7	682	NISP_LACIA	Q07596 lactococcus

34	170.5	5.7	688	1	LIP_STAEP	Q02510 staphylococ
35	170.5	5.7	970	1	PSU1_YEAST	P53550 saccharomyc
36	170.5	5.7	1260	1	ALSI_CANAL	P46580 candida alb
37	170.5	5.7	1435	1	LTEL_YEAST	P07866 saccharomyc
38	170.5	5.7	1636	1	BUD3_YEAST	P25558 saccharomyc
39	170	5.6	1803	1	YUL3_YEAST	P47024 saccharomyc
40	169.5	5.6	571	1	TACY_STRPY	P21131 streptococc
41	169.5	5.6	837	1	CSG_HAIWO	P25062 halobacteri
42	169.5	5.6	1020	1	BCA_STRAG	Q02182 streptococ
43	169.5	5.6	2859	1	RBP1_PLAVB	Q00798 plasmodium
44	169	5.6	731	1	BAF1_YEAST	P14164 saccharomyc
45	169	5.6	954	1	BIR1_YEAST	P47134 saccharomyc

## ALIGNMENTS

RESULT 1	FNBA_STAU	STANDARD;	PRT;	1018 AA.
ID	FNBA_STAU			
AC	P14738;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Fibronectin-binding protein precursor (FNBP).			
GN	FNBA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NCITC 8325-4;			
RX	MEDLINE=89098998; PubMed=2521391;			
RA	Signaes C., Raucsi G., Joensson K., Lindgren P.-E.,			
RA	Anantharamiah G.M., Hoeck M., Lindberg M.;			
RT	"Nucleotide sequence of the gene for a fibronectin-binding protein			
RT	from Staphylococcus aureus: use of this peptide sequence in the			
RT	synthesis of biologically active peptides.;			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).			
CC	-1- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN			
CC	PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE			
CC	WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO			
CC	THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S. AUREUS,			
CC	THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE			
CC	INVASION.			
CC	-1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by			
CC	an amide bond (Potential).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib.ch">license@sib.ch</a> )			
CC	-----			
DR	EMBL; J04151; AAA26632.1; -			
DR	InterPro; IPR004237; Fn_bind.			
DR	InterPro; IPR001899; Gram_pos_anchor.			
DR	Pfam; PF00746; Gram_pos_anchor; 1.			
DR	Pfam; PF02986; Fn_bind; 1.			
DR	TIGRFAMS; TIGR01167; LpxTG_anchor; 1.			
DR	TIGRFAMS; TIGR01168; YSIRK_signal; 1.			
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.			
KW	Cell wall; Peptidoglycan-anchor; Repeat; Signal.			
FT	SIGNAL	1	36	
FT	CHAIN	37	985	FIBRONECTIN-BINDING PROTEIN.
FT	PROPEP	986	1018	REMOVED BY SORTASE (POTENTIAL).
FT	REPEAT	545	574	B-1.
FT	REPEAT	575	604	B-2.
FT	DOMAIN	745	878	4 X APPROXIMATE TANDEN REPEATS,
FT				FIBRONECTIN-BINDING DOMAIN.
FT	REPEAT	745	782	D-1.

FT REPEAT 783 820 D-2.  
 FT REPEAT 821 859 D-3.  
 FT REPEAT 860 878 D-4 (INCOMPLETE).  
 FT DOMAIN 879 948 5 X TANDEM REPEATS, PRO-RICH (WR).  
 FT REPEAT 879 892 WR1.  
 FT REPEAT 893 906 WR1.  
 FT REPEAT 907 920 WR3.  
 FT REPEAT 921 934 WR3.  
 FT REPEAT 935 948 WR3.  
 FT SITE 982 986 LPTG SORTING SIGNAL (POTENTIAL).  
 FT MOD\_RES 985 985 AMIDE-LINKED TO CELL WALL (POTENTIAL).  
 SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;

Query Match 10.0%; Score 300.5; DB 1; Length 1018;  
 Best Local Similarity 21.2%; Pred. No. 6.8e-06;  
 Matches 123; Conservative 118; Mismatches 243; Indels .97; Gaps 25;

35 GIKRSEDRETTNTNDENATFLQKTPDQNT---THLTVEEVKSSSVSSSIDTAQ 90  
 30 GQDEKAASQKTTVEEN-----GNSATDNKTSQTOTATVNHLEETSYNATV--TE 82  
 QY 91 QPSHTT-INREE--SVQSDNVEDSHVDFANSKIKESNTESGKEENTIEOPNKVEDS 146  
 83 QPBNATQVTEEARPKAVQAQTAQNPANIEVKEVKE-----EAKPVKETT 130  
 QY 147 TTSQPSG--YTNIDEKISODELLNPINEYENKARPLSTSNOPSIRVT-----V 196  
 131 QSDNSGDDQROVDLTPKKAQNOVAETQV---EVAQPRASESKPRVTSADVAEAKEA 186  
 QY 197 NOLAEQSGVNVHLIKYTDQSITEGVDSGVKAHAEMLIYDVEVDKYSQDGTMT 256  
 187 SNAKVEGTIVTSKYVEIGSI--EGHNNINKV--EPHAGQRAVLKYKLFEGHLOGGYFD 244  
 QY 257 VDIKNTVPSPDLTDFSTPIRKDMS-----GEIATGYDNKNQOITVETTYDYDKYENI 311  
 245 FTLSNNVHTGVSFARKVPEIKGSVYMATGEVLEGS-----KIVTETNOIEDKVDV 297  
 Db 312 KAHKLTSLYDKSKVPNNNTKLDVEYKTAASSV-----KTIYEYORPNDRTANLQSM 366  
 298 TAELEMLFLDPTKTVQNGN-----OTITSLNEQTSKELDYKYKYGNGYVYANLNGS 351  
 QY 367 FTNIDFNHTVEQTIYINPLRYSAKETNVNISGDEGSLIIDSTLIKKYK--VGDQNL 425  
 352 IEFNKANNRFSHAFKFP--NNGKTSVYVTGTLMKGSQNGNQPVRFELIGNEDI 409  
 Y 426 PDSNRIV---DYSEYEDVYNDYDAQLG--NNNDVINFINGIDSPYLIKYSKYPKND- 478  
 410 AKS--VYANTDTDSKFEVTSNMSGNLNONGSYSLNIELDKTYVYVHDEGLNGTDE 467  
 QY 479 -DITTIQOYVMTQTTINEYIGEFRTASYDNTIAFSTSSGQCGDLP-----EKTYSIGD 532  
 Db 468 VDEFT--QMGVGHPEQLYKYRGTTLWDNGLVLYSNKANGNEKNGPILIONNKEFYK--- 523  
 QY 533 YVMEVDKQGIQNTNDNEKPLSNVLTLYPDTGSKSVRTD 573  
 Db 524 ----EDTIKETLIGQYDKN-----LVTTVEEYDSSTLIDID 555

RESULT 2  
 ARP\_PLAFA STANDARD; PRT; 537 AA.  
 AC P04931;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Asparagine-rich protein (AG319) (ARP) (Fragment).  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=66176787; PubMed=2421257;  
 RA Stahl H.-D., Bianco A.E., Crewther P.E., Burkot T., Coppel R.L.,

RA Brown G.V., Anders R.F., Kemp D.J.:  
 RT "An asparagine-rich protein from blood stages of Plasmodium  
 RL falciparum shares determinants with sporozoites.";  
 CC Nucleic Acids Res. 14:3089-3102(1986).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.ch](mailto:license@isb.ch)).

DR EMBL; M24328; AAA29491.1; -  
 DR PIR; A23770; A23770.

FT NON\_TER 1 1  
 FT NON\_TER 537 537  
 SQ SEQUENCE 537 AA; 63000 MW; A6D76B15318CC239 CRC64;

Query Match 6.7%; Score 203; DB 1; Length 537;  
 Best Local Similarity 17.8%; Pred. No. 0.037;  
 Matches 93; Conservative 106; Mismatches 212; Indels 112; Gaps 19;

QY 11 NNQSTINTDNNQIIEKEETNNYDGIKRESDRETTNTNDENATFLQKTPDQNTLITE 70  
 Db 4 NNNNNNNNDGDN--INVQNTNEKDNKNNFKQYNNNFKPEENNNNSNTMSRNSNVE 61  
 QY 71 EEVKESSSVSSNSSID--TAQPSHTT-----NREESVQ--SDNVEDSHVDFAN 119  
 62 EHLR-ANSLIMNNSNTINNTYNOOTRSPFMENENENKKNYHFGGMNNNHFKKYYNNN 120  
 QY 120 SKIKESNTESGKEENTIEOPNKVEDSTSPSGYTNIDEKISODELLNPINEYENK 178  
 121 SSMK--NTDNNKTDTSYMKGTINNDN-----NMODYLRINIINNEYKGS 163  
 Db 179 ARPLSTSAQPSIKRYVNDLAEQSGVNVHLIKYTDQSITEGVDSGVKAHAEMLI 238  
 164 AKKKEFTYN-----YMKNNLKLFTQNNNDNMNINEDNNNNNNNNNGVFSNQNNNMN 217  
 QY 239 YDVFEDVDKVGSGDTVTVIDKNTVPSPDLTDFSTPIRKDNGSLIATGYDKNKQIT 298  
 218 RNNNSINKRLN-----NNNNNNNNNNKKSQKQNSNNNFYNNVYQARKSMN 268  
 QY 299 YTFDVKYENIKAHKLTSLYDKSKVPNNNTKLDVEYKTAASSV----- 345  
 269 NNN 328  
 QY 346 -KTIYEYORPNDRTANLQSM-----FTNIDFKNTV-----EDTIYIN----- 384  
 Db 329 VKASMMYNN--NESVTANPQOMNEQNTNDNMKRENNNNNNYGYDDVTVHNNNTPTSDP 387  
 QY 385 -----PLRYSAKETVNI-----SCNGD-----EGSTIIDSTLIKKYK-----VG 420  
 388 FSRAGVGYNNYLNNNNNNSVYNNSSNGNNMKRENSKKNVAAQNNQSLNNKNNNNNNIN 447  
 Db 448 MNEISNNNTLNNNNEYNQNNNEDDDDDGELGEDKXIDIN 490

RESULT 3  
 PI4K\_DICDI STANDARD; PRT; 1093 AA.  
 AC P54677;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Phosphatidylinositol 4-kinase (EC 2.7.1.67) (PI4-kinase) (PtdIns-4-  
 DE kinase) (PI4k-alpha).  
 GN PI4K OR PI4K  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.

OX NCBI\_TaxId=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AX3;  
 RX MEDLINE=96009592; PubMed=7565716;  
 RA Zhou K., Takegawa K., Emr S.D., Firtel R.A.;  
 RT "A phosphatidylinositol (PI) kinase gene family in Dictyostellum  
 discoidium: biological roles of putative mammalian p110 and yeast  
 Vps34p PI 3-kinase homologs during growth and development."; *Mol. Cell Biol.* 15:5645-5656(1995).  
 RL Mol. Cell Biol. 15:5645-5656(1995).  
 CC -1- FUNCTION: ACTS ON PHOSPHATIDYLINOSITOL (PI) IN THE FIRST  
 CC COMMITTED STEP IN THE PRODUCTION OF THE SECOND MESSENGER  
 CC INOSITOL-1,4,5-TRISPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1L-myo-inositol = ADP +  
 CC 1-phosphatidyl-1D-myo-inositol 4-phosphate.  
 CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U23479; AAA85725.1; -.  
 CC DictyDb: DD01102; P1KD.  
 CC InterPro: IPR000403; P13\_P14\_Kinase.  
 CC Pfam: PF00454; P13\_P14\_Kinase.1.  
 DR SMART, SM00146; P13Kc.1.  
 DR PROSITE: PS00915; P13\_4\_KINASE.1; 1.  
 DR PROSITE: PS00916; P13\_4\_KINASE.2; 1.  
 DR PROSITE: PS50290; P13\_4\_KINASE.3; 1.  
 DR Transferrase; Kinase; Multigene family.  
 FT DOMAIN 146 156 ASP-RICH.  
 FT 184 200 POLY-ASN.  
 FT 203 206 POLY-ASN.  
 FT 227 233 POLY-ASN.  
 FT 277 315 POLY-ASN.  
 FT 441 445 POLY-ASP.  
 FT 445 466 POLY-THR.  
 FT 494 501 POLY-GLY.  
 FT 501 501 POLY-THR.  
 FT 696 696 POLY-GLN.  
 FT 761 772 POLY-THR.  
 FT 775 785 POLY-THR.  
 FT 833 1093 P13K/P14K.  
 SQ SEQUENCE 1093 AA; 123017 MW; 8762BC78355AA635 CRC64;  
 Query Match 6.7%; Score 202; DB 1; Length 1093;  
 Best Local Similarity 20.0%; Pred. No. 0.093;  
 Matches 116; Conservative 100; Mismatches 223; Indels 142; Gaps 24;

OY 266 SDLTDSFTIKIKDNGS---EIIATGYDNKKNKQITFTFDYVDKYNKAKHLKLTSYID 322  
 DB 469 NHLKSTTSGVGINSSSTPININSAGAGAGGGGGINH-----GYDI-----SYLD 515  
 OY 323 KSKVPNNNTKL---DVEYKTA-----LSYVKRTIVE---VORPENTNANQS 365  
 DB 516 KCKTPPESKLSHDHFEFLSKRCDYLDILSFIOKLHISILPLPIDLQAKLKH 575  
 OY 366 MFTNIDKRNHTVEQTIYINPLRSAKETNVNISGNGEGSTIIDSTTIKYKVGDNQNL 425  
 DB 576 ---EISLILNPLGLYV-PLMQSS-----NHHGVRI-----P 605  
 OY 426 PDSNRIDYSE-----YEDVTNDYAOLGNNNDVNIFFGNDSPYIIKYSKYDPKDD 479  
 DB 606 PEEVKILNSRRRPFLLVLEIESEHALSN-----LFEVVSF-----LQ 647  
 OY 480 YTFIQGTVMQTINEYTGE---FRTASVNTIAFSTSS 515  
 DB 648 YTGNSALKKDDIRKRYSEKFKSFINSINSTSNSSDS 688  
 RESULT 4  
 HMW1\_MYCGE STANDARD; PRT; 1139 AA.  
 ID HMW1\_MYCGE  
 AC Q49413; Q49365;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cytochrome high molecular weight protein 1 (Cytochrome accessory  
 DE protein 1).  
 GN HMW1 OR MG312.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxId=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischman J.L., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Luster J.C.,  
 RA Peterson S.N., Smith B.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium.";  
 RT Science 270:397-403(1995).  
 RL [2]  
 RP SEQUENCE OF 721-847 FROM N.A.  
 RC STRAIN-ATCC 33530 / G-37;  
 RX MEDLINE=94075230; PubMed=8253680;  
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;  
 RT "A survey of the Mycoplasma genitalium genome by using random  
 RT sequencing.";  
 RL J. Bacteriol. 175:7918-7930(1993).  
 CC -1- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH  
 CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS  
 CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW  
 CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHESIN PROTEINS  
 CC IN THE MYCOPLASMAL MEMBRANE AT THE ATTACHMENT ORGANELLE (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: LOCALIZES SPECIFICALLY TO THE ATTACHMENT  
 CC MEMBRANE (BY SIMILARITY).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U39712; AAC71534.1; -.  
 CC EMBL: U02261; AAD12527.1; -.

DR TIGR: MG312: -  
 KW Cytochrome: Structural protein: Complete proteome.  
 SQ SEQUENCE 1139 AA; 13051 MW; 0011D3288C3D0836 CRC64;

Query Match 6.5%; Score 196.5; DB 1; Length 1139;  
 Best Local Similarity 18.0%; Pred. No. 0.17;  
 Matches 150; Conservative 116; Mismatches 276; Indels 293; Gaps 34;

QY 1 SSDEKADVNNNOSTITDNNQI--IKKE-----ETNNYDGIKRSDDRT 44  
 DB 343 SSSFEETPTPEOFOVNEVNDQFKEPEVLESFKNQODVETSDLSNENLSENN 402  
 QY 45 ESTNNVENATFLQ-----KTPQDNH-----LTSEVK 74  
 DB 403 KDATNNDLSNSETIQLNSSETASDVHYSKSEPIHDYKFGSOLSQSNNSNLSSEPVK 462  
 QY 75 ESSSV-----ESSNSSID-----TAQPSHTTIRRE----- 101  
 DB 463 FNSETAPDAHFESESEVVDQVYIYQNEELKPTLDQPSDDYAKQPTDENIGFDDLP 522  
 QY 102 -SVQTSNVEDSHVDFANSKIKESNTSGKEENTIEOPNKVKEDSTTSOPS----- 152  
 DB 523 PEYKQPSVVDQPSDDHFAKQPESTDTSYSPDSDLP-----TLQPSLDDHYQYN 575  
 QY 153 -----GTNDEKISQDEL-----LNLP 171  
 DB 576 FDHHEELKPAEENQNYVGFDOVQANLDNNEELQPTAEKVTTFESKQAVVDSYQLP 635  
 QY 172 IN-EYENKAPPLSTTSQPSIKRVTYNQLAAEAGSNVNHLLK--VTQSTIEGVDSEGV 228  
 DB 636 IDTQDQGTFFSSSFQPTVEQD-----QVNSEVNDQFKEPTKEPEVLESFKNQDV 689  
 QY 229 IKAHDAENLIYDVTEVDKY-----KSGDTMTVDIDKN 262  
 DB 690 VETSNTYNNLOKQEDIQSDNKITITTKKSSQPIPTLPISFVSNIEKRPVETTLADKES 749  
 QY 263 -----TVPSDLTDSFTPK-----IKNSGEIATG--TYDNKKNQITTFDYVD 306  
 DB 750 QQEQITNSTEDSKTLAKTLVOLQOINLSNOSIVTSSEVLKDDQDQTLTNTVSED 809  
 QY 307 KYENIKAKLITSYIDSKVPNNNTKLDVEKKTALSSVKN-----TTVEYQRPENRTA 361  
 DB 810 QQRIEYFVFAKEVEHEISTON--KQSVDEKSELDNFKKSKDLKIISELKRELNP 867  
 QY 362 NLSMFTNIDTKNHTVQGTIYNPL-----RYSAKETVNI 397  
 DB 868 NFDATFQNMNYQMSVKOSFIHLNDFVNTYKQISERYLLIKKELQSELSRLIDQENMLNV 927  
 QY 398 SGNDSESTIIDSTIIR-----YKVGDN-QNLPDSKRIYD----- 434  
 DB 928 QFNNAKMLTLQKEEMRSLASDAFALVKBPSNYSLOKSGEIMRHQORAITENKIES 987  
 QY 435 -----SEYEDVTNDYQALQNN--NDVNINFGNIDSPYIIKIVISKYD-----PNK 477  
 DB 988 IQGSLQKLQTYVNSCCETIMNNIKKLDNTLFAKKEKDPPL--LSNDSVTNDGLVEPQ 1045  
 QY 478 --DDYTIQGTVMQTTINEYTGFEFRASVDNTAESTSSQGGQDLPEKTYKIGDYVW 535  
 DB 1046 LMDL--LIDSNMFPDINSNEQDLDFIYEMMDRNIIDFEFEGFN-----NDFV- 1089  
 QY 536 EDVQ--KDIGQNNDNEKPLSNVLTLYPDGTSK--SVRDED-----GKYQFD 581  
 DB 1090 -DIDAKVMSKSAFVNDLLETLY-----PDRTSNFSLLDEDLFESSGDFSLD 1138

## RESULT 5

YMDB\_CAEEL STANDARD; PRT; 918 AA.  
 AC P34487;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical 96.6 kDa protein f59B2.12 in chromosome III.

GN F59B2.12  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-B15C01 N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Harkins T., Hillier L., Jier M.,  
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Steden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,  
 RA Wohldman P.;  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans";  
 RL Nature 368:32-38(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z11505; CAAT7581.1; -  
 CC PIR: S31132; S31132.  
 CC WormPep: F59B2.12; CE01024.  
 DR WormPep: F59B2.12; CE01024.  
 KW Hypothetical protein.  
 SQ SEQUENCE 918 AA; 96560 MW; E464FD86B1495DE CRC64;

Query Match 6.4%; Score 192; DB 1; Length 918;  
 Best Local Similarity 21.2%; Pred. No. 0.2;  
 Matches 138; Conservative 90; Mismatches 278; Indels 146; Gaps 26;

QY 4 EKKDVLNNNOSTITDNNQIKKKEETNNYD-----GIKRSSEDRK-----STN 49  
 DB 113 ESKVDEANENKIKISAGS-VIEGKSHNKSDDASSYGLKSSKTYADKGMILSSTN 171  
 QY 50 VDENEA-----TFLOKTPQDNTHL-----FEEVKESSVSSSSIDTAQOP 92  
 DB 172 KINNQSSALDEGNEFYNOQADGTFURNNTGKKNIDESHVNLDENAQMSIG-ADGT 230  
 QY 93 SHTTINREESVQTSN-VEDSHVDFANSKIKESNTSGKEENTIEOPNKVKEDSTTSOP 151  
 DB 231 SHNITNRKGVSGDHNMAADAH-----SNFESLDAQNKKSSQNYSKKAASAG- 278  
 QY 152 SGYINIDEKISNOBELNLPINEYENKARPLSTS-----QPSIKRVTYN-----OLA 200  
 DB 279 --SNADFE-SNLESLKNAADGTSMSNSTGNENNTSYDKATAEVMSKKNVADGTSMEA 334  
 QY 201 AEQSNVNHLLKYVDQSTTEGYDSEGVYKADHAENLIYDVTEFVDDQVKKSGDTPTVID 260  
 DB 335 SHAGSNSSKINSASGQSSDLSMGPNG-IKSHSTSNKTDN--YALDENQAGSISSEIG 391  
 QY 261 K-----NTVPSDLTDSFTIIPKINDNGEIIATGTYNKKNQIYTT 300  
 DB 392 KNGCRSLNESSIESGRKAESRNNTAADTLDSVDANGVTSSSHSKSASGTSIDENHNKTHA 451  
 QY 301 FTDYVDKYENIKAKLITSYIDSKVPNNNTKLDVEYKTA--LSSVN-----KTTTV 350  
 DB 452 LQASVDEHGNKKNHSDIGSYRNKKTGEGNSEMASISINAGDTSOVAVXKNDTRNRYEA 511  
 QY 351 EYQRPENRFTANLQSMFTNIDTKNHTVQGTIYNPLRYSAKETNVNISGNDSESTIID 410  
 DB 512 EKSALERNHENSQGTCKD-----ESKGSNRVN-RIDGGSNLAVG 551

FT	CARBOHYD	817	817	(POTENTIAL)
FT	CARBOHYD	838	838	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	841	841	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1018	1018	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1082	1082	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1100	1100	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1103	1103	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1113	1113	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1200	1200	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1571	1571	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	CARBOHYD	1593	1593	N-LINKED (GLCNAC . ) (POTENTIAL)
FT	DOMAIN	95	106	POLY-GLN.
FT	DOMAIN	283	286	POLY-ASN.
FT	DOMAIN	1283	1292	POLY-HIS.
FT	DOMAIN	1651	1660	POLY-GLN.
SQ	SEQUENCE	1664 AA:	187859 MW; BA2EFDDC8196790 CRC64;	
<hr/>				
Query Match 6.3%; Score 189.5; DB 1; Length 1664;				
Best Local Similarity 20.4%; Pred No. 0.51; 236; Indels 171; Gaps 31;				
Matches 131; Conservative 105; Mismatches				
<hr/>				
QY	8 DVINNNOS--INTDDNNQIIRK-----EETNNYDGIEKRSEDRTSTYT 48	:	: ::	:
Db	138 DKVENENHAPTYINTSPNKSIKKATPKASPPKVAFVTYNPELHNHF--DNRYEEEDOQQ 195	::	:: ::	:
QY	49 NVDENEAETFOCKTPODNTHTLEEVKESSSVESNSSIDTAOOPSHTT-----INRE 101	:	: ::	:
Db	196 KEDSVEPLIQ-----HOKKDPSQFNVSDERTWASVPT--PRLHTTKPTFPALLNKN 247	:	: ::	:
QY	102 SVQTSNDVEDSHVD-----FAASKIKE-----SITESGKEENTIEDQPKVKVEDSTT 148	:	: ::	:
Db	248 EV---NSEPALTDMLKRFENSLSDEKVNILYSTNNNSKNVSDMSHDLO---N 299	:	: ::	:
QY	149 SOPSGYTNIDEKISNODELMLPINEYENKRAPL-STPSAOPSIKFYVNOIAEOGSNV 207	:	: ::	:
Db	300 LQDSAKNTENNINHLSALKAPKPDITEN---PLNSLTNAOISLASGSSQSLOSRLRD 356	:	: ::	:
QY	208 NHLIKV---TDQSITEGYDSEGVYKAHADENLIVDTFEVDDKYVSGDJMTVIDDKNTV 264	:	: ::	:
Db	357 NRVLSEVPGPSKPKNVPGSLNDGIKGFSB-----EVVESILL 392	:	: ::	:
QY	265 PSDLT-DSETLPKID----NSGGILTNGTYDNKNKOITYTFMDYVDKEYENIKAH---- 314	:	: ::	:
Db	393 PRDLRSRDLKETTEKHDAPEHHNNEFDIAKSTFNKNG-OVLVSDDHLSPDSQSYNHTEQS 451	:	: ::	:
QY	315 -LKLSYSYDKSKVPMN-----NTKLDEVYKTALSVMK 346	:	: ::	:
Db	452 ILNLNLSAQSOGISLALEKORQTOEQOTAABEEETSFSFNDIKVAOEPRSNIJEPKV 511	:	: ::	:
QY	347 TTIVEYQRPNERNANLASMFTNIDTKNH-TVEQTIIYNPLRYSAKETNVNISG----- 399	:	: ::	:
Db	512 TIKKEPVATEIKAKREFESSRIIRIKINEDIELAEPAADHPKKKENANSIHVEDCTDALUKKA 571	:	: ::	:
QY	400 -NGDGSGTIIDSTLIIX-YKVGDONLIPDSNRITDYSEYEUVTVNDUYAOQLNN---NDV 454	:	: ::	:
Db	572 LNDEDSOPTQOSTMWSRFRHIDSOWMKLEDSDND---GDRED--INDDISREKSOLDIANG 625	:	: ::	:
QY	455 N-----INFGNIDSPFIYIKISKYPDPKDDYTTLIOCVTMQTTLINEYTGGER-----T 502	:	: ::	:
Db	626 SOTSIDIIGDKIGNSSSEITTKTLLA---PPSRONNDKENKSLSDPADRNANNESSLOOQLEVPHRT 682	:	: ::	:
QY	503 ASYDNTIAFSTSQGOGGDLPEEK---TYKIGDY-VWEVDVK 540	:	: ::	:
Db	683 KEDDSILANSSNIA-----PREELTLPVLEANDYSSFNDYRK 719	:	: ::	:
<hr/>				
RESULT 7	LIPA_MYCPU	STANDARD:	PRT;	578 AA.
ID AC Q50274; Q98039;	LIPA_MYCPU	STANDARD:	PRT;	578 AA.
DT 16-OCT-2001 (Rel. 40, Created)				
DT 16-OCT-2001 (Rel. 40, Last sequence update)				
DT 15-JUN-2002 (Rel. 41, Last annotation update)				

```

Yy 174 EYEKAPRLPTSTASQPSIKRTVYVQQLAEQSGSVNHLIKV-----TDQSIT--EEYD 223
Db 194 NRQKK-----NEVKDDKKALLQOMQKMLNESASILESESYDQTSLSLTFKESMP 240
Oy 224 -DSEGVIRKHADEANLIVDTFVEVDKYSQDPTMTVDIDKNTVPSPD--LTDSPITPIKDN 280
Db 241 LGLEVLVK---LEMLDSHEEKEISFKTTNGVQVWLLTSSNLTSGKMKIKSFSDKTYSH 297
Oy 281 SGELIATGYDNKKK---QITTYFTDYVDKYENIKAKLKL-TSYIDK---SKVPNNNTKL 333
Db 298 SPTEETTFDEFETNKELKQERIEKIOKAIKDYQ-----IKIKONYKDKPLISKVALSNFDL 352
Oy 334 DVEKTKLSSVKNKTIITYEYQRPENRRANLQSMFT-----NIDKKNH-----TWQGT 380
Db 353 NLNRKD-LEIVNNSLKFDGSLNODLSEKQKITFEKVSENOAKTKRAHKITTTLDNL 411
Oy 381 IYINPLRSAKETVWNISGNDGEGSTIID--STIIKYK--VGDNONLPPDS---NRI- 431
Db 412 VFEKTLKMSYK-TKKEYLDEFKNSALMDDLOALSTVFEKSLMHPIYLPRAKSKININT 470
Oy 432 -----YDYSEYEDVTNDYVQALGNNDVYNINEGNIDSPYIIVISKYDNPKNKY 480
Db 471 LINDVASFCQGYDYL-----DDF-----NGSAKLKF-----KLQGEQROBIT 508
Oy 481 TTIOQTYMTQTTIMEYGEFRTASYDNTIARTSSGQSGQDLPPE--KTYIIGDVWEDV 538
Db 509 FTINGFLKVSLLIDPLKGNLRNSEFD-VKASSNGYMLQOYTTAAEVFNHYSNGKSYWAT 567
Oy 539 DKDG 542
Db 568 ANDG 571

RESULT 8
EBAL_PLAFC STANDARD; PRT: 1435 AA.
ID EBAL_PLAFC
AC P19214;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Erythrocyte-binding antigen EBA-175.
OS Plasmodium falciparum (Isolate Camp / Malaysia).
OC Eukaryota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90377299; PubMed=2204835;
RA Slim B. K.L.;
RT "Sequence conservation of a functional domain of erythrocyte binding antigen 175 in Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 41:293-296(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on ways CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb.ch/announce/ or send an email to license@isb.slb.ch).
CC -----
Cc CC
DR EMBL; X52524; CA36756.1; -.
DR PIR; S11561; S11561.
KW Antigen.
FT DOMAIN 159 1104
FT ERYTHROCYTES.
FT VARIANT 1031 1031 E->V (IN STRAINS FOR-3 AND ITG).
FT SEQUENCE 1435 AA; 167388 MW; 32A4309021B1C3D6 CR64;
Oy Query Match 6.3%; Score 188.5; DB 1; Length 1435;
Best Local Similarity 20.5%; Pred. No. 0.47;
Matches 114; Conservative 102; Mismatches 199; Indels 141; Gaps 28;
1 S D S E D K N D V I N N G S I N T D N N Q I I K K E T N Y N D C ----- I E K R S E D R I E S T T N V D E N 53

```

```

Db 875 NSDSEEEVYVNH-----ISKPSINNGDDSGSGSATVSSSSSSNGSLDIDDRN 923
Qy 54 EATFLQRTPODNTLITEEV--KSSSVSSSSSIDTAQPSHTTINEESVQTSNDVED 111
Db 924 GDFEVRTPQDNTANT-----EDVIRKENADKDEKGADEERHSTSELSPEEKMLTDNCG 979
Qy 112 SHVDFANSKIKE--SNTESGKEEVTIOPKKVKE-DSTTSOPSGTYNIDE----- 159
Db 980 NSLN---HEEVKEHTNSNDNQOGGIVNMVVEKELKDLENPS--SSIDEKKAHELSE 1034
Qy 160 HINODELMLPINEYENKARPLSTSAOPSIKRVTVNOLAEGSGVMNLIKTVDOSI 218
Db 1035 PNLSSDDMSKTP-----GPLDNTS-EETTERISNNEKYNV-----REDERTL 1077
Qy 219 TEGYDSEGVIAKIH-----DAENLIYDTEFEVDKVKSGDTMTVIDKNTVPSDLTDSF 272
Db 1078 TKYEDI--VLKSHNRESDDGE--LYD-----ENSLDSTYNDESEDAEAKMGND 1124
Qy 273 TIPKIDNSGEIATGYTDKNKKQITTYFDYVDKYEINIKAKLITSTIDSKYVNNNTK 332
Db 1125 T-SEMSSHSSOHISSDOCKNMKTYGDLGTHVON---EISVPYTGTEIDEXLRESKESK 1179
Qy 333 LDVEYKTLSSVNTTITVEYQRPENRFTANLOSM--FTNIDTKNTVEQITINPLRYSA 390
Db 1180 IHKADEERLSTHD---IHKINP-EDRNSNTLHLKDINREEVERHLTQNTINISQERDLO 1234
Qy 391 K---ETANNISGNGDEGSTIIDSTITIKYVKGDNQNLPSNRIDYSEYEDVTNDYAO 447
Db 1235 KHGHTIMNHLGDC-----VSEKQINSHSHGNQD-----R 1266
Qy 448 LGNNNDV-----NINSGNIDSPYTIKYISKYD-----PNKDYTTIQOTVMQTTI 493
Db 1267 GGNNGVNLNMRSSNNNNENNIPSRY-----NLDDKLLDLLENRNDST-KELIKKLAEI 1320
Qy 494 NEYGEFTASYDNTI 509
Db 1321 NKCENISVKYCDHMI 1336

RESULT 9
LMA2_HUMAN STANDARD: PRT: 3110 AA.
AC P24043; Q14736; Q93022;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy
chain).
GN LAMA2 OR LAMM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94124633; PubMed=8294519;
RA Vuolteenaho R., Nissinen M., Salola K., Byers E., Eddy R.,
RT Vuolteenaho H., Shows T.B., Sariola H., Engvall M., Tryggvason K.,
RT "Human laminin M chain (merosin): complete primary structure,
RT chromosomal assignment, and expression of the M and A chain in human
RT fetal tissues.";
RL J. Cell Biol. 124:381-394 (1994).

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97066955; PubMed=8910357;
RA Zhang X., Vuolteenaho R., Tryggvason K.,
RT "Structure of the human laminin alpha2-chain which is
RT affected in congenital muscular dystrophy.";
RL J. Biol. Chem. 271:27664-27669 (1996).

RN [3]
RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.

```

```

RC TISSUE=Placenta; PubMed=2185464;
RX MEDLINE=90238994;
RA Ehlig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.;
RT "Merosin, a tissue-specific basement membrane protein, is a
RT laminin-like protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268 (1990).

RN [4]
RP VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.
RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
RA Marzluft G.A., Amato A.A., Mendell J.R.;
RT "Novel single base polymorphisms and rare sequence variants in
RT the laminin 2-chain coding region detected by RNA/SSCP analysis.";
RL Hum. Mutat. 13:174-174 (1999).

RN [5]
RP ERRATUM.
RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
RA Marzluft G.A., Amato A.A., Mendell J.R.;
RL Hum. Mutat. 13:340-340 (1999).

CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end.
CC -1- THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
CC 4 (S-MEROSIN).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: PLACENTA, STRIATED MUSCLE, PERIPHERAL NERVE,
CC CARDIAC MUSCLE, PANCREAS, LUNG, SPLEEN, KIDNEY, ADRENAL GLAND,
CC SKIN, TESTIS, MENINGES, CHOROID PLEXUS, AND SOME OTHER REGIONS OF
CC THE BRAIN; NOT IN LIVER, THYMUS AND BONE.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -1- DISEASE: DEFECTS IN LAMA2 ARE THE CAUSE OF MEROSIN-DEFICIENT
CC CONGENITAL MUSCULAR DYSTROPHY (MCMD).
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC -1- THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -1-
DR EMBL: 226553; CAA81394.1; -
DR EMBL: U66796; AAB18388.1; -
DR EMBL: U66733; AAB18388.1; JOINED.
DR EMBL: U66734; AAB18388.1; JOINED.
DR EMBL: U66735; AAB18388.1; JOINED.
DR EMBL: U66736; AAB18388.1; JOINED.
DR EMBL: U66737; AAB18388.1; JOINED.
DR EMBL: U66738; AAB18388.1; JOINED.
DR EMBL: U66739; AAB18388.1; JOINED.
DR EMBL: U66740; AAB18388.1; JOINED.
DR EMBL: U66741; AAB18388.1; JOINED.
DR EMBL: U66742; AAB18388.1; JOINED.
DR EMBL: U66743; AAB18388.1; JOINED.
DR EMBL: U66745; AAB18388.1; JOINED.
DR EMBL: U66746; AAB18388.1; JOINED.
DR EMBL: U66747; AAB18388.1; JOINED.
DR EMBL: U66748; AAB18388.1; JOINED.
DR EMBL: U66749; AAB18388.1; JOINED.
DR EMBL: U66750; AAB18388.1; JOINED.
DR EMBL: U66751; AAB18388.1; JOINED.

```





```

Db 2314 G-LMNRERKES-----DCKGCTVSPQVDESDGCTQFDG 2345

RESULT 10
TANA_XENLA
ID TANA_XENLA STANDARD; PRT: 1744 AA.
AC 001550;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanabiti.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=radpole head.
RX MEDLINE=92398961; PubMed=1524825;
RA Hammati-Britvanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate
RT neurons defines a new class of intermediate filament protein."
RL Neuron 9:417-428(1992).
CC -1- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURULA AND PERSIST
CC DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
CC CORD.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M99387; AAA49966.1;
DR PIR; JH0720; JH0720.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament_1.
DR PROSITE; PS00226; IF_1.
DR Intermediate filament; Coiled coil; Neurone.
FT DOMAIN 1 12 HEAD.
FT DOMAIN 13 314 ROD.
FT DOMAIN 315 1744 TAIL.
FT DOMAIN 8 48 COIL 1A.
FT DOMAIN 49 60 LINKER 1.
FT DOMAIN 61 156 COIL 1B.
FT DOMAIN 157 179 LINKER 12.
FT DOMAIN 180 193 COIL 2A.
FT DOMAIN 194 199 LINKER 2.
FT DOMAIN 200 314 COIL 2B.
SQ SEQUENCE 1744 AA; 199561 MW; 6502EAC9F66CAE93 CRC64;

Query Match 6.1%; Score 184.5; DB 1; Length 1744;
Best Local Similarity 19.7%; Pred. No. 0.86;
Matches 149; Conservative 122; Mismatches 239; Indels 245; Gaps 42;

OY 1 SDEKKNVNNNSINTDNNQIKKETNN--YDGIEKRSDERTSTTNDENATFL 58
DB 886 SFOIKDFEINQECLEKSPQIREAPTEVDHQVDFMQGSEFEREVGLNNIKQEVDTL 945
OY 59 QKTPD-----NTHLIEEVK--ESSV-----ESSNSIDF- 88
DB 946 QNYDEDSFONNDEPOLESCLDQEQKIKLEENQJSENGNONGFNGNDIEEFSQOQYDID 1005
OY 89 -----AQQPSHTINRE-----SVOTSDNVEDSHVDFANSKIKESN 126
DB 1006 EICQETIGNQVSAQLCESDINQDKLSMEDDEEQNNPTEEDNIGLEQESQENTRENEGT 1065

```

```

OY 127 TESGKEENTIEOPNKVEDSTTSQPSGYNTIDKISNO--DELLNLPINE--YENKARPL 182
DB 1066 KFSQECDCVYKRP-----ED--MSDKSEYSGQOEDLDKQVDTSLNDQANNLLEKEEYIL 1119
OY 183 STTSAPSI-----KRVYNQLAEGSGSVNHLIKVTDQSI--TEGY----- 222
DB 1120 HHADQRSVNDIEITIDKLSERIIDNELAT--VDVNSLANKQVDFLDEYAVDDV 1176
OY 223 ---DDESEGYKAK-----DAENLYDYTFE-----YDKKSGDPTWYDI 259
DB 1177 GMQDDDSQGYQTKEDLFVDGNNIIKIEIQSTLSLNOEICERVDWYDEDI--SGEAKNESV 1235
OY 260 DKNTPV-----SDTDS---FTPIKQNSGRIANGTYDNKNKQITTF 301
DB 1236 ENNDVVDVLPKAVKVTGDEQISPLQDEKLNLEMTEDTKDNDGOL---CLEKNE----- 1285
OY 302 TDYVDKVENIK-----AHLKTSYIDSKVY--NNNKL-----DVEYKALSS-- 343
DB 1286 TEYIEVTSPOCATLSDHAGREL--VDQNSANLQFCENPRTKTLAHHIEYETVADSIL 1343
OY 344 -----VNKTTVEYQRPENFTANLQSMFTNIDFKNHTVEQTIYINPLYSAKETNVI 397
DB 1344 ESTEBOVQETERIPF--KPEDSKMEN-----ENSESESVDSQESISLN---SHKSEEFEEI 1393
OY 398 SGNQDEGSIIDDSITLIKVKYKQDNQNLPSDKRIYD--SEYEDVYN--DDYAGLQNNNDY 454
DB 1394 S-----KDYOL--EQLPDPVPLPLEDEFEDELTEQDPVHEHONNDS 1435
OY 455 NINFGNIDSPYIIKYSKYDPNKKDYTTTQQTVMQTTINEYT--GEFRTASYDNNTAFST 513
DB 1436 GA-----STFTTSV-----DEDKEREYRESVSDNESENEEPFGDVLSDVTSQVEYVT 1483
OY 514 SSGQGGDLPEPKTYKIGDYVEDVYDKGIQNTND--NEKPLSNVL----- 557
DB 1484 LSLGLAQ-----PPSY-LGD---NESESDMENAEILNENPSDIYDFVNSQMTETKIITA 1534
OY 558 -----VTLVYPDQTSK---SVRTDEGKXQFD 581
DB 1535 EQVTEQTEVTLOFDDAPNKLNTLNAREKETDYDE 1569

RESULT 11
YGA4_YEAST
ID YGA4_YEAST STANDARD; PRT: 817 AA.
AC P46949;
DR 01-NOV-1995 (Rel. 32, Created)
DR 01-NOV-1995 (Rel. 32, Last sequence update)
DR 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 90.8 kDa protein in RRP41-SNG1 intergenic region.
GN YGR196C OR G7589.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RC MEDLINE=96076633; PubMed=7502584;
RA Guerrero I., Mala e Silva A., Barreiros T., Arroyo J.,
RA Garcia-Gonzalez M., Garcia-Saez M.I., Rodrigues Pousada C.,
RA Nomela C.;
RT "The complete sequence of a 9000 bp fragment of the right arm of
RT Saccharomyces cerevisiae chromosome VII contains four previously
RT unknown open reading frames."
RL Yeast 11:1087-1091(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: X82775; CA58019.1; -
DR EMBL: Z72981; CA97222.1; -
DR SGD: 50003428; YGR196C.
KM Hypothetical protein.
SQ SEQUENCE 817 AA; 90797 MW; E52C5D659D63BEBB CRC64;

Query Match 6.1%; Score 184; DB 1; Length 817;
Best Local Similarity 21.0%; Pred. No. 0.37;
Matches 134; Conservative 88; Mismatches 237; Indels 180; Gaps 31;

QY 1 SDEPKNDVNN-----NOSINFDNNNOI--IKKEFNNDGIEKSED 42
DB 120 SSEHSKSDYLSTASLSPSENKSPHTRNRAVNDLNLQISREMPF---IKQSD 176
QY 43 RTESTTAVDE--NEATFLQKTPQDNTLHTEEVKSSSVSSNSIDTAQPSHTTINRE 100
DB 177 RRDSDSCDEIENAPLGEAVPSSSPVEDEKSHSLGVSMDTNEADTF--NPTTRGN 233
QY 101 ESQVTSQNVEDSHVSDPANSKIKESN-----TESGKEENTIEQPNKVKEDSTISQPSGYT 155
DB 234 EHLSSDGVSEQKDEP---KVSERGLADILPAKEEN-LQQ---EDGVEVSSGAL 284
QY 156 NIDKISNDELNLPINEYENKARPLSTSAQPSI---KRVYNQOLAEOGSVNNHLIK 212
DB 285 EKKESEKTSIRN-----RN-----STSGQDKVAKPKVAVNETKSDNGYRNSPFND 333
QY 213 VTDSITEGYDSEGVAKAHDAENLIYDVTFPEVDKYSQDNTVIDDKNTVPSDLTD-- 270
DB 334 YQHSDEEDDNNNEGSSSD-----DDNRSSVADKHADIRQSKQJLDTDD 381
QY 271 --STTIPIKIKNSGEIITATGYDNKKQIYTFDYDKYIKAHKLITIDKSKVYN 328
DB 382 ALSTY-ESIKYSTETEEDEEDNESIE-----DKNE-----DNESIED 419
QY 329 NNTLDEYKATLSVNTIYEVQRPEN-----RTANLQSMFTNIDKN 374
DB 420 ENEDTD-SYKFSNREKSGILTSDEEEKMSDSDSGSKAPSGYFSKMIG-DDKG 477
QY 375 HTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDSTIIKYKGNQNLPSNRIYDY 434
DB 478 DSAIQPNQIDIT---ENTNLNSGSELENSDDEEDHINCKYLESSVAKDSTVDW 533
QY 435 SEYEDVTMDYAOQGNNDVNIINFGNIDSP--YII-----KVISKYD-- 474
DB 534 KPDSALRSGFQ-----DTANKKAPPGYVIDSNGKLVDLTPASMKRIVYSTSEM 584
QY 475 -----PNK---DDYTIQOTVPMQTTINETYTGFEFRASYDNTIAFSTSGQG--QGDLP 523
DB 585 ESTMDAFPSKEDDDLETITRTKIT-----YDNNITIYVPGILGNQMLP 629
QY 524 P-----EKTYRIGDYVWEDVDKGIQNTNDRKPLSNVL 557
DB 630 PLPMDAQEQNLNAGN-----DNSTNDNSNNTANDL 660

RESULT 12
RBP2_PLAVB STANDARD; PRT; 1251 AA.
AC 000799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9231338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax

```

```

RT merozoites";
RL Cell 69:1213-1226(1992).
CC -FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: M88096; AAA29744.1; -
KW Malaria; Receptor; Membrane.
FT NON_TER 1
FT NON_TER 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 6.1%; Score 184; DB 1; Length 1251;
Best Local Similarity 20.0%; Pred. No. 0.61;
Matches 140; Conservative 115; Mismatches 236; Indels 210; Gaps 37;

QY 11 NNNOSINTDDNNQIJKKEE--TNNYDGIKRSERDTE-----STNV--DENATFLQK 60
DB 302 SNSKNVNIENINENIRNSEOYLKIDEAKQASTKVELFKHETTISIFKESELIGVET 361
QY 61 TPQDNTLHTEEVKSSSVSSNSIDTAQPSHTTINREESVQTSQNVEDSHVSD-FAN 119
DB 362 KSQKIKKAKEDIMKE---IERHNEIQOVAKGPOENLKNLKNPHNYCAEDELNDRKTN 418
QY 120 SKIK-ESNTESGKEENTIEQPNKVKEDSTISQPSGYTNIDKISNDELNLPINEYENK 178
DB 419 AKVILETLESVK--HNLSFTINIKQG-----EKIYSK-----AKDIMQK 457
QY 179 ARPSTISQAQSIKRVYNQOLAEOGSVNNHLIKYTDS---IPE-----GYDGS---E 226
DB 458 IKATSENTEAETLEKV---KDDQSNVYVNLQTTETERNLIVTEKNLNGIDSTYIE 512
QY 227 GVIRKAHAENLIYDVF-EVDKYSQDNTVIDDKNTVP-----SDLTDSFTPIKIKD 279
DB 513 GALKESKGN---YEIGFLEKLEIKGNKKLKVYDIKKKINSTGVNFSLPFNFDLQYDF 569
QY 280 NSGEIATGYDNKKQIYTFDYDKY-ENIK-----AHLKLTSTYDQSKV 326
DB 570 NKN-----INDYENKKEGIYNEFEGSLNKISENLRASENSTSDYNSAKTLRLAOKREYNL 625
QY 327 PN-----NNTKLDEP-----YKTALSSVNTIYEVQRPENTANLQSMFTNID 371
DB 626 LNKEEANKYLRDVAKKVESFRIFENMKESLDKINEMIKKEQLTVNEGH-GNKKQLVENIK 684
QY 372 T-----KNHTVEQTIYINPLRYSAKETNVNISGNGDEGSTIIDSTIIK 415
DB 685 ELVDENNLSDILKQATGNNEI-QKITHSTLKNKKK---TILGHVDSAKVVG---IK 735
QY 416 V-----YKVGNNQNLPSNRIYDYSEYEDVTNDYAOQGNNDVNIINFGNIDS 463
DB 736 ITPELATLLELDGAKLKAQELKPESKNNVVLEFENMSK-----NTNELDVH-KNIOD 787
QY 464 PY--IIVKISKYD---PNKDDYTIQ--QIVTMQTTINETYTGFEFR----- 502
DB 788 AKVALETLIHAHSEIDTWKQDSKSLIEGNQYILKVVLLNOKKNKISSIKSKEEAVSVKI 847
QY 503 -----ASYDNTIAF---STSGQCGQGLDPEKTYKIGDYWEDVDKGI 543
DB 848 GNVSKKHSELKTIQSDSYNIILAEKOTELQNLNRSFTQKLT----- 891
QY 544 QNTNDNEK-----PLSNVLTLYPDGTSKSVRTDED 575
DB 892 -NTNDSKLEIKITFDESLKNALKTL---EGEVNALKASSD 928

```



```

CC -1- PFM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND 5.8S GENES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X16561; CAA34560.1; -.
CC PIR: S07485; RNZQ2L.
CC InterPro: IPR000684; RNA_polII_repeat.
CC InterPro: IPR000272; RNA_pol_A.
CC InterPro: IPR002879; RNA_pol_A2.
CC Pfam: PF00623; RNA_pol_A; 1.
CC Pfam: PF01854; RNA_pol_A2; 1.
CC PROSITE: PS00115; RNA_POL_II_REPEAT; 9.
CC Transferase: DNA-directed RNA polymerase; transcription; zinc; Repeat;
CC KM DNA-binding; Nuclear protein; phosphorylation; zinc-finger.
CC FT ZNFING 68 84
CC FT DNA_BIND 378 411
CC FT ZNFING 707 725
CC FT DOMAIN 1093 1128
CC FT DOMAIN 1144 1159
CC FT DNA_BIND 1182 1193
CC FT DOMAIN 1687 1694
CC FT DOMAIN 1258 1290
CC FT DOMAIN 1261 1290
CC FT DOMAIN 1602 1612
CC FT DOMAIN 1746 1759
CC FT DOMAIN 1806 1820
CC FT DOMAIN 2061 2246
CC FT DOMAIN 2247 2384
CC FT SEQUENCE 2452 AA; 278166 MW; F995E117617A48F CRC64;

Query Match 6.1%; Score 183; DB 1; Length 2452;
Best Local Similarity 21.1%; Pred. No. 1.5; Mismatches 275; Indels 172; Gaps 28;
Matches 145; Conservative 95;

3 DEERNDVYNNQSIINTDNNQIIKKEETNNYDGIKESSEDTTESTTVDNEATFLOK-- 60
Db 1283 DDDLDDEENYSNDINIGNRKYNGTILKNND-----ENSLMNPIDVHKVNNFLEKLV 1336
QY 61 -TPQDNLHLEEEVKESSSVSSSIDTAAQPSHTTINREESVQTS-----DN 108
Db 1337 IIRKINSN-----DRLSYEAQNNATILKAHLRTYLNLSKLQTRHVSYKGLDWLQE 1389
QY 109 VE-----DSHVSDFANSKIKESNTES-----GKEENTTECP-----NKV 142
Db 1390 IEKIFYISLCPGRCVGLAAQSIGEPATQTLTFPHFAGVGSNNVLYGVRLEKELINIV 1449
QY 143 KEDSTQPSGTYNIDEXISNDELNLPINEYENKARPLSTSAOPSIKRVYNOQLAAE 202
Db 1450 KNVNT---PSTTIYLDQVNSVDOOKADILKLETTIKQLTSHAQIIYDNTTTTILEE 1506
QY 203 QGSNVNHLIKYTDOSTIEGYDSSGVIRKAHDAENLIYDVFVDKVSQDTMTVIDIKN 262
Db 1507 DKSVMNEPEEPEDDDTQ-YSLGEMVLR-----IQLTINHVNEK-KLTKELEYIYIS 1537
QY 263 TVPDDLDLSTFIKIKONSSEII-----ATGYDNKNKQITTYTFDYVKNENIKAH 315
Db 1558 VFSSDELD---IYTDNSEDLVLRIRVKYLNGEYNEFMMNDYVDVNAEVOQDEDEEHL 1614
QY 316 -----KLTSYIDKSKVPNN-----NTR-----LDVEYKTAIS 342
Db 1615 VANDRGVDETCKNSTHPIHHDYNNMTNIFKSKVKNNISSDINTKNEDSISINSSNNEQV 1674

```

```

QY 343 SYNKITVEYQRPENENTANLQSMFTIDKKNHVE-----QTIYNPLRYS 389
Db 1675 NISSPVSNNMHHNNNNNNNDSSNINDIKYKNIKKEGNCALRGGDSNTSALFGKNS 1734
QY 390 AKETNV--NISGNDGSESTIIDSTIKYKVGDNQNLPSNRIRYDYSEVEDYTNDD-- 444
Db 1735 QKEDNIYNNNDNDND--DEEEEDLFGDHANSPKNTKGKKNKNNKNNNNENK 1790
QY 445 YAOGLNNNDVNNF--GNIDSPYIIRVSKYDPKDDYTTIQYTMQTTINEYGEFR 501
Db 1791 NKSNNNSNNNSNTYDDGDVDN-----DNDNDNDNKSQDI-----TIKE----- 1829
QY 502 TASYDNITAF-STSSGCGGGLDPEKTYKIGDYWEVDYDKDQIONT--NDNEKPLSNVL 557
Db 1830 ---DNDVAEMKSTKNAEEDL---ELKNNIENHISREDTEDETLKLMEOCLSTLK 1881
QY 558 -----VLTYPDGTSKSVRTDEDKY 578
Db 1882 LRGIENITKYMRRESKITYDSNGKF 1908

RESULT 15
YCG1_YEAST STANDARD; PRT: 853 AA.
AC P25588; P25589; P27513; P87003;
DT 01-MAY-1992 (Rel. 22, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 97.9 kDa protein in CHA1-KRRI intergenic region.
GN YCLO61C OR YCLO61C/YCLO6C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid:4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Rasmussen S.W., von Wettstein D.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RL [2]
RP REVISIONS.
RA Gromadka R.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 384-596 FROM N.A.
RC STRAIN=ATCC 28383 / F1100;
RX MEDLINE=90384830; PubMed=2169608;
RA Kern L.;
RL "The URA1 gene of Saccharomyces cerevisiae encoding uridine kinase.";
RL Nucleic Acids Res. 18:5279-5279(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59720; CAA42405.1; -.
CC EMBL: X53998; -. NOT_ANNOTATED_CDS.
CC PIR: S19391; S19391.
CC PIR: S19391; S19391.
CC PIR: S19392; S19392.
CC PIR: S21360; S21360.
CC PIR: S29373; S29373.
CC SGD: S0000566; YCLO61C.
CC Hypothetical protein.
KM CONFLICT 505 505 L -> V (IN REF. 3).
FT CONFLICT 567 567 MISSING (IN REF. 3).
SO SEQUENCE 853 AA; 97946 MW; 16E09FCC0BF248D1 CRC64;

Query Match 6.1%; Score 182.5; DB 1; Length 853;
Best Local Similarity 18.9%; Pred. No. 0.45;

```

Matches 121; Conservative 117; Mismatches 234; Indels 169; Gaps 29;

```
QY 5 EKNDVNNNSINTDNNQ---IKKEETNNYDGIKRSERDRETTNVDENEATFLQK 60
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 291 EENNFOLNHSGSDSGSGFALSGNEIADYESSGSENNRRSDSEKEDDEIILKQK 350
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 TPQDNTHL-----TEEEVKESSSVESNSSIDTAQPSHTTN---REBSVOTSDNV 109
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 351 KSHVKKHINESDSDTEVEAKPKKEADES-----LPKRINMIGHYGNIGEDTDKF 402
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 110 FDSHVSOFPANSKIKESNTESKENTIEQPNKVEDSTSCPGYTNIDEKISNQ---DE 166
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 403 QETVVLDTQNTL-----EVAERNTLE--NEVKDQVYNEEA-----DEAIRQLIDKE 449
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 LLLNLPINEYENKARPLSTSAOPSIKRYTVN--QLAEQGSNNHLIKVTQDITGEYD 224
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 450 KLOLKOKEKEHEAK-----IKELKKRGVTNFEKEAESEDEWHGIGADGEGSDYDS 503
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 225 -----SEGVTKADHAENLIYDVFEDDKYKSGDTMTVDID-----KNIVPS 266
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 504 DLEKMDIDYSKNNFNPHEIREMLAENKEMD--IKMINKIIVYDIKNGGFRRNKRANKSL 561
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 267 DLTSFTIPKIKDSGEIATGTYDNKKQITFTFDYVDKXENIKAHKLKLSYIDKSKV 326
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 562 ELSD-----DDEDDVLQ--QYRLKRRELMR-----KRRLEIG--DDAKL 596
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 PNNNTKLDVEYKKTALSSVNTKITVEYORP---NENRTANLCSMTNIDTKNHTVEQTIYI 383
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 597 V-KNPKSAFFESMVEDI-----IFKNPFGAEFEYNLDITATDLDT----- 639
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 384 NPLRYSAKETVNIISGN-GDEGSTIID--STIKVYKVGDNQNLPSNRIDYSEYED 439
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 640 -----ODNSINVGDNNTGNNEQKPVQKNNKKVLIISEDVCKSLSLKSN---NYEDFE 688
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 440 VTNDYALQGNNDVNIINGNIDSPYIIKVISKYPNKNDDYTTIOQVIMQTTINEYGE 499
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 689 -TDKELSRIOHGND EAI-----EDLYTLKONSSIKSFTNSQT-D 725
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 500 FRTASVDNTIAFSTSSGOGGDLPEPKTYKIGDYVMEVDKDIQNTNDNE----- 550
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 726 STTSKTVNTIIDLEKRPEDDEVEENGDSLYGVFKHPSTIKSFASTRDINDKFKGKNTV 785
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 551 -----KPLSNVLVTLTYPDGTSKSV---RTDEDGKYQFD 581
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 786 KILKSYKTVGSSKASITVYMGKTRKLIAPKRKTEGSHRYHHD 826
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 14, 2003, 12:59:00  
Job time : 19.6487 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 12:56:59 ; Search time 22.2509 Seconds

(without alignments)  
5389.420 Million cell updates/sec

Title: us-09-147-405b-13

Perfect score: 3013  
Sequence: 1 SSDEKNDVYNNNSINTDD.....PDGTSKSVRTDEDGKYQFDG 582

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3013	100.0	1092	2	070022 staphylococ
2	2810.5	93.3	931	2	09KI13 staphylococ
3	1160.5	38.5	1171	2	09KW6 staphylococ
4	1144.5	38.0	1166	2	086489 staphylococ
5	1137.5	37.8	1141	16	099W46 staphylococ
6	1137.5	37.8	1141	16	0932F7 staphylococ
7	576	19.1	1315	2	086488 staphylococ
8	559	18.6	1385	16	099W47 staphylococ
9	499	16.6	1733	2	09KI14 staphylococ
10	489	16.2	953	16	099W48 staphylococ
11	460.5	15.3	881	2	093KH7 staphylococ
12	455	15.1	947	2	086487 staphylococ
13	452	15.0	933	2	053653 staphylococ
14	452	15.0	935	16	0932C5 staphylococ
15	452	15.0	989	16	099V14 staphylococ
16	397.5	13.2	877	16	099R07 staphylococ

17	396	13.1	913	2	086476 staphylococ
18	376	12.5	940	2	053682 staphylococ
19	344.5	11.4	961	16	099RD3 staphylococ
20	326	10.8	1038	16	099RD2 staphylococ
21	268.5	8.9	1301	5	08WKS5 plasmodium
22	266.5	8.8	1160	2	08RM86 streptococ
23	266	8.8	2402	2	09AER7 staphylococ
24	261	8.7	1161	2	09X3M7 streptococ
25	259	8.6	1161	2	08RJ10 streptococ
26	239	7.9	1039	2	P72534 streptococ
27	237.5	7.9	2276	2	093TY6 staphylococ
28	234	7.8	604	5	026021 plasmodium
29	231.5	7.7	1401	2	08RJR4 staphylococ
30	228	7.6	2647	5	0904X0 plasmodium
31	225	7.5	1463	2	086919 staphylococ
32	220	7.3	1843	5	0963L5 dictyostell
33	219.5	7.3	1364	5	08T223 dictyostell
34	217.5	7.2	559	5	0903Y8 plasmodium
35	216.5	7.2	1959	16	08RYN1 listeria mo
36	216.5	7.2	3394	5	077384 plasmodium
37	215.5	7.2	1564	2	093IM3 amyasaeta moo
38	215	7.1	518	12	09DH49 mycoplasma
39	214.5	7.1	2081	10	09LH98 mycoplasma
40	212.5	7.1	1125	16	098PM9 mycoplasma
41	212.5	7.1	3130	5	09BK46 plasmodium
42	212	7.0	2771	5	026216 plasmodium
43	211.5	7.0	1344	2	049545 mycoplasma
44	211.5	7.0	2206	5	096205 plasmodium
45	211	7.0	691	5	09GSD1 plasmodium

## ALIGNMENTS

### RESULT 1

ID 070022

PRELIMINARY:

PRT: 1092 AA.

AC 070022;  
DT 01-AUG-1998 (TREMblrel. 07, Created)  
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Fibritogen-binding protein precursor.  
OS Staphylococcus epidermidis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Staphylococcus;  
OX NCBI\_TaxID=1282;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HB;  
RA MEDLINE=98261511; Pubmed=9596732;  
RX Nilsson M., Frykberg L., Flock J.I., Pel L., Lindberg M., Guss B.;  
RT "A Fibritogen-binding protein of Staphylococcus epidermidis."  
RL Infect. Immun. 66:2666-2673(1998).  
DR EMBL: Y17116; CA76638.1; -;  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
DR TIGRPFAMS: TIGR01167; LPXTG\_anchor; 1.  
DR TIGRPFAMS: TIGR01168; YSIRK\_signal; 1.  
DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
KW Signal.  
FT SIGNAL 1 51  
FT CHAIN 52 1092  
SQ SEQUENCE 1092 AA; 119292 MW; 6542BC39ADBB984 CRC64;  
Query Match 100.0%; Score 3013; DB 2; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 1.9e-105;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SSDEKNDVYNNNSINTDDNNQIKKETNNYDGIKRSDDRESTTNVDENKATFLQK 60  
DB 75 SSDEKNDVYNNNSINTDDNNQIKKETNNYDGIKRSDDRESTTNVDENKATFLQK 134  
QY 61 TPQDTHLTLEEVKSSSVESNSSIDTAQPSHTTINRESVQISDVAVEDSHVSDPFANS 120

```

Db 135 TPQDNHLEETEEVKESSSVSSNSSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANS 194
QY 121 KIESNTESGKEENTIEQPKVKVEDSTSPSGYINDEKISNODELNLPINEYENKAR 180
Db 195 KIESNTESGKEENTIEQPKVKVEDSTSPSGYINDEKISNODELNLPINEYENKAR 254
QY 181 PLSTSAOPSIKRVYVNOQLAEOGSVNVNHLIKYTDOSITBEGYDSEGVKKAHAENLIYD 240
Db 255 PLSTSAOPSIKRVYVNOQLAEOGSVNVNHLIKYTDOSITBEGYDSEGVKKAHAENLIYD 314
QY 241 VPEEVDKVKSGDPTMVDIDKNTVPSDLTDSFTPIKIDNSGELIATGYDNKNKOITYT 300
Db 315 VPEEVDKVKSGDPTMVDIDKNTVPSDLTDSFTPIKIDNSGELIATGYDNKNKOITYT 374
QY 301 FTDYVVKYENIKAKHLKTSYIDSKVPPNNNTKLDVEYKKTALSSVNTTIVEYORPENRT 360
Db 375 FTDYVVKYENIKAKHLKTSYIDSKVPPNNNTKLDVEYKKTALSSVNTTIVEYORPENRT 434
QY 361 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGEGSTIIDSTIIKYYKVG 420
Db 435 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGEGSTIIDSTIIKYYKVG 494
QY 421 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIIKVISKYDPNKDXY 480
Db 495 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIIKVISKYDPNKDXY 554
QY 481 TTIOQVTVMTQTTINEXYTGERTASYDNTIAFSTSSGOGGDLPEPKTYKIGDYVWEDVK 540
Db 555 TTIOQVTVMTQTTINEXYTGERTASYDNTIAFSTSSGOGGDLPEPKTYKIGDYVWEDVK 614
QY 541 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 582
Db 615 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 656

```

## RESULT 2

```

O9K113 PRELIMINARY; PRT: 931 AA.
ID 09K113
AC 09K113:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fibrinogen-binding protein sdrG.
GN SDRG.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K28;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eldhin D., Lina G.,
RA Speziale P., Foster T.J., Hook M.;
RT "The serine-aspartate repeat (sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245042; AAF72510.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRfams: TIGR01167; LPTXG_anchor; 1.
DR TIGRfams: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN; 1.
SQ SEQUENCE 931 AA: 102955 MW: 591E657D97027116 CRC64:

```

## Query Match

Best Local Similarity 93.3%; Score 2810.5; DB 2; Length 931;  
Matches 545; Conservative 15; Mismatches 21; Indels 1; Gaps 1;

```

QY 1 SDEEKNDVNNOSITDNNQIIRKEETNNNDIGIKRSDEDTSTTNDENATFLQK 60
Db 74 SSNEKNDVNNOSITDNNQIIRKEETNNNDIGIKRSDEDTSTTNDENATFLQK 132

```

```

QY 61 TPQDNHLEETEEVKESSSVSSNSSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANS 120
Db 133 TPQDNHLEETEEVKESSSVSSNSSIDTAQOPSHTTINREESVQTSNDVEDSHVDFANS 192
QY 121 KIESNTESGKEENTIEQPKVKVEDSTSPSGYINDEKISNODELNLPINEYENKAR 180
Db 193 KIESNTESGKEENTIEQPKVKVEDSTSPSGYINDEKISNODELNLPINEYENKAR 252
QY 181 PLSTSAOPSIKRVYVNOQLAEOGSVNVNHLIKYTDOSITBEGYDSEGVKKAHAENLIYD 240
Db 253 PLSTSAOPSIKRVYVNOQLAEOGSVNVNHLIKYTDOSITBEGYDSEGVKKAHAENLIYD 312
QY 241 VPEEVDKVKSGDPTMVDIDKNTVPSDLTDSFTPIKIDNSGELIATGYDNKNKOITYT 300
Db 313 VPEEVDKVKSGDPTMVDIDKNTVPSDLTDSFTPIKIDNSGELIATGYDNKNKOITYT 372
QY 301 FTDYVVKYENIKAKHLKTSYIDSKVPPNNNTKLDVEYKKTALSSVNTTIVEYORPENRT 360
Db 373 FTDYVVKYENIKAKHLKTSYIDSKVPPNNNTKLDVEYKKTALSSVNTTIVEYORPENRT 432
QY 361 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGEGSTIIDSTIIKYYKVG 420
Db 433 ANQSMFTNIDTKNHTVEQTIYINPLRYSAKETNVNINSGNGEGSTIIDSTIIKYYKVG 492
QY 421 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIIKVISKYDPNKDXY 480
Db 493 DNQNLPSNRITYDYSEEDVTNDYVQALGNNNDVNIINFGNIDSPYIIKVISKYDPNKDXY 552
QY 481 TTIOQVTVMTQTTINEXYTGERTASYDNTIAFSTSSGOGGDLPEPKTYKIGDYVWEDVK 540
Db 553 TTIOQVTVMTQTTINEXYTGERTASYDNTIAFSTSSGOGGDLPEPKTYKIGDYVWEDVK 612
QY 541 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 582
Db 613 DGIQNTNDNEKPLSNVLTLYTPDGTSKSVRTBEDGKYOPDG 654

```

## RESULT 3

```

O9KMX6 PRELIMINARY; PRT: 1171 AA.
ID 09KMX6
AC 09KMX6:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bone sialoprotein-binding protein.
GN BBP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=024;
RX MEDLINE=20115096; PubMed=10642520;
RA Tung H.S., Guss B., Hellman U., Persson L., Rubin K., Ryden C.;
RT "A bone sialoprotein-binding protein from Staphylococcus aureus: a
RT member of the staphylococcal sdr family ]."
RL Biochem. J. 345:611-619(2000).
DR EMBL: Y18653; CAB57332.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRfams: TIGR01167; LPTXG_anchor; 1.
DR TIGRfams: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN; 1.
SQ SEQUENCE 1171 AA: 127123 MW: C5BC812F9DASAB84 CRC64:

```

## Query Match

Best Local Similarity 38.5%; Score 1160.5; DB 2; Length 1171;  
Matches 271; Conservative 99; Mismatches 182; Indels 79; Gaps 20;

```

QY 1 SDEEKNDVNNOSITDNNQIIRKEETNNNDIGIKRSDEDTSTTNDENATFLQK 51
Db 1171 SDEEKNDVNNOSITDNNQIIRKEETNNNDIGIKRSDEDTSTTNDENATFLQK 1171

```



```

Db 57 STENAKODEASADNKEVSETEENNSTOKNDLTPIKKETNTDSHOEAKKAPTSSVQOQ 116
QY 52 ENEATFLQKTPQDNTHLLEEVEKES---SSVESSNSSIDTAQOPSHTTINEESVQTSN 108
Db 117 ONNATSTETKPPQ--TEKENVKPSTKTATEDTSVLEEKKAPNNT--NNDVTTKPS-- 170
QY 109 VEDSHVSDFNASKIKESNT---ESGKEPNTIEOPNKVED---STTSQPSQTYNI--DEK 160
Db 171 -----TSEIQTPTTPOESTNIENSOPPTPSKVNDQVDTDTNPKPPVANSKEEL 220
QY 161 TSNDELLNLPINE--YENKARPLSTSAOPSIKRV-----TYNQLAAEGSNVNLHILK 212
Db 221 KNNPEKLEKELVRNDSNTDRSKPAPATAPTSAAPKRVNAKIRFAVAQAPRAAVASNNVNDLIT 280
QY 213 VTQDSITEGYDSEGVAKAHDAENLIYDVFEEVDKYSKSGDTMTVDIDKNTVPSDLTDSF 272
Db 281 VTQKITEGICKD--DGVIOAHGHEHIIYTSDEKIDNNAVAGDPTMTVKYDKETIPSDITDDE 339
QY 273 TIRIKRNSGELIATGYDKNKNOITFTFYDVKYENIKAHKLTSYIDSKVPNNNTK 332
Db 340 TPVDITDPSGEVIAKGFDLNTKTYKFTDYVDRYEVNNAKLELNSYIDKKEVP--NETN 398
QY 333 LDVEYKTALESNKITIEYORPENRPTANLOSMTNIDTKNHTVEQTIYINPLRSYAKE 392
Db 399 LNTFFATADKETSNAKVEYKPIYKDESNISQISFSLDTKHEVEQTIYINPLKNAKN 458
QY 393 TNVNI-----SGND---EGSTIIDSTIIKVKYKVDNQLPDSNRIDYSEYEDVND 443
Db 459 TNVTKSGVANDGYTGDSTIIDSTIISNTEIKYKVASGQOLPOSNKLYDSQYEDVYNS 518
QY 444 DYADLGNN---NDVNINGNINDSPIIIVISKYDPNKDDTYTIOQTYMOTINETYEEF 500
Db 519 --VTINKNYGTMANINENGIDISAIYKRVSKYTPGABDDLAVOOGVMTT---TNKY 571
QY 501 RTASYDN-----TAFSTSSGOGGDL--PPEKTYKIGYVWEDVDKDIQNTDNNEK 551
Db 572 NYSSTAGTILFYQLTLVTVTSVK-----PEEKLYKIGYVWEDVDKVGQGDTSKEX 626
QY 552 PLSNVLVLTYPDGTSKSVRTDEDKYOPDG 582
Db 627 PMANVLVLTYPDGTSTKSVRTDANGHYEFGG 657

RESULT 4
ID 086489 PRELIMINARY; PRT: 1166 AA.
AC 086489;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Sdr E protein.
N SDR E.
JS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_TaxID=1280;
[1]
RP SEQUENCE FROM N.A.
GN STRAIN-NEWMAN.
RC MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCrea K., NI Elidhah D., O'Connell D., Cox J., Hook M.,
RA Foster T.J.;
RA "three new members of the serine-aspartate repeat protein multigene
RT family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL; AJ005647; CA06652.1;
DR InterPro; IPR001899; Gram_Pos_anchor.
DR Pfam; PF00746; Gram_Pos_anchor.1.
DR TIGRFAMs; TIGR01167; LpYTG_anchor.1.
DR TIGRFAMs; TIGR01168; YsIRK_signal.1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
SQ SEQUENCE 1166 AA; 126548 MW; 750A7B0135287D4A CRC64;

```

```

Query Match 38.0%; Score 1144.5; DB 2; Length 1166;
Best Local Similarity 42.6%; Pred. No. 1.4e-35;
Matches 263; Conservative 108; Mismatches 191; Indels 55; Gaps 18;

QY 1 SSDEKNDVNNNOSITDDNNQI--IKKETNNYDGIKESDRTESTVNDENEATFL 58
Db 66 ATTDNKNKVEYSETEENNSTOKNDLTPIKKETNTDSHOEAKKAPTSSVQOQNNVTAT 123
QY 59 QKTPQDNTHLLEEVEKES---SSVESSNSSIDTAQOPSHTTINEESVQTSNVEDSHVS 115
Db 124 TETPQDN--TEKENVKPSTKTATEDTSVLEEKKAPNNT--NNDVTTKPS-----TS 172
QY 116 DFANSKIKESNT---ESGKEPNTIEOPNKVEDSTTSQPSQTYNI--DEK 171
Db 173 EPSTSEIQTPTTPOESTNIENSOPPTPSKVND--NOYDATNPKPPVANSKEELNPN 229
QY 172 -----INEYENKARPLSTSAOPSIKRV-----TYNQLAAEGSNVNLHILKYTDQ 216
Db 230 EKLKELVRNDSNTDRSKPAPATAPTSAAPKRVNAKIRFAVAQAPRAAVASNNVNDLITK 289
QY 217 SITEGYDSEGVAKAHDAENLIYDVFEEVDKYSKSGDTMTVDIDKNTVPSDLTDSF 276
Db 290 TIKVG--DGKDVAAAHGDKDIEYDTEFTIDKRVKKGDTMTINYDKNVPSPDLTKNDPID 348
QY 277 IKDMSGELIATGYDKNKNOITFTFYDVKYENIKAHKLTSYIDSKVPNNNTKLDVE 336
Db 349 ITDPSGEVIAKGFDLNTKTYKFTDYVDRYEVNNAKLELNSYIDKKEVP--NETSLT 407
QY 337 YKTALESNKITIEYORPENRPTANLOSMTNIDTKNHTVEQTIYINPLRSYAKETNVN 396
Db 408 FATAGKETSQAVTYDQOPMVAHGSNLSITFKLDEKQTIIEQOLYVPLKKSMTNFVVD 467
QY 397 ISGNG-----DEGSTIIDSTIIKVKYKVDNQLPDSNRIDYSEYEDVND--DYA 446
Db 468 IAGSQVDYDGNIKLNGSTIIDSTIIDONTIEIKYVNSQOOLPOSNRIYDFSQYEDVYSQEP 527
QY 447 QLGNNNDVNINGNINDSPIIIVISKYDPNKDDTYTIOQTYMOTINETYGEPTASYD 506
Db 528 KFSNNVATLDPGDINSAYIIRVSKYTPSTDGELDIQGSMTT--DKY-GIYVAGYS 585
QY 507 NTIAFSTSSGOGGDL--PPEKTYKIGYVWEDVDKDIQNTDNNEKPLSNVLVLTYPDG 565
Db 586 NFIVTSNDTGGDGTVRKEEKLYKIGDYVWEDVDKDYOGGDSKEXPMANVLVLTYPDG 645
QY 566 TSKSVRTDEDKYOPDG 582
Db 646 TTKSVRTDANGHYEFGG 662

RESULT 5
ID 099W46 PRELIMINARY; PRT: 1141 AA.
AC 099W46;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
N protein.
OS SDR E SA0521.
OC Staphylococcus aureus (strain N315).
CC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_TaxID=158879;
[1]
RP SEQUENCE FROM N.A.
GN MEDLINE=21311952; PubMed=11418146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Takasumaru H., Maruyama A., Murakami H., Hoshoyama A., Mizutani-Oi Y.,
RA Taira H., Kihara S., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Oshima K., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shibata T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;

```

CC bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

```

RESULT 7
086488
ID 086488 PRELIMINARY; PRT: 1315 AA.
AC 086488;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Sdrd protein.
GN SDRD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RX MEDLINE=99098700; PubMed=9884231;
RA Josefsson E., McCreia K., Ni Eldin D., O'Connell D., Cox J., Hook M.,
  Foster T.J.;
RT "Three new members of the serine-aspartate repeat protein multigene
  family of Staphylococcus aureus."
RL Microbiology 144:3387-3395(1998).
DR EMBL: AJ005646; CA06651.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS: TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
SQ SEQUENCE 1315 AA; 14275 MW; 58DAE1F48E6A689 CRC64;

Query Match 19.1%; Score 576; DB 2; Length 1315;
Best Local Similarity 30.1%; Pred. No. 2,8e-14;
Matches 186; Conservative 112; Mismatches 228; Indels 92; Gaps 29;

QY 1 SDEEKNV---INNOSTIDDDNNQIIKKEETNNNDGIEKREDETESTNVDEN-NEAT 56
DB 55 STNKELENTTASDNOSSDKVDMQOLNODNTRKNOKEMVSSQGNETTSGNKLIEE 114
QY 57 FLOKTPQDNTLHEEVEKSSSVSSSSSIDTQAQPSHTTINREESVQTSNVEDSHVD 116
DB 115 SVOSTGNKVEVSTAKSDQASPKSTNEDLNTRQ-----TISNOELQ-PDLGKNSYVN 168
QY 117 FANSKIKESNTESGKEENTIEOPNKKEDSTTSQPSGYNTI-DEKISNOELL-NLPIN 173
DB 169 -----VQPTNEBKRVKVD-----AKTESTT-----LNKSDAISKNDDETLVDNNSNS 209
QY 174 EYENKARPLSTTSAP-----SIKRVVNOLAAEOGSNVNLIKVTQDSIEGVDSGVI 229
DB 210 NNEENADITLPKSTAPKRLTRKRIAAVOPSSTEAKNVNDLITSNTTLTYVVDADKNNKIY 269
QY 230 KAHDAENLIYDVFEEVDKVKSGDTMTVDIDKNTV-----PSDLTDSFTIPRIKD-NSG 282
DB 270 PQDYLSLKSQIT--VDKVKSGDYFTIKY-SDTYQVYGLNPEIDIKN--IGDIKPNNG 323
QY 283 EIIATGTYNKKKQITFTFDYDKYENIKAHKLTSYIDSKSVPPNNKTLDEYKTAIS 342
DB 324 ETIATKAKHDTANLLITFTFDYVDRENSVOMGINSIYMDADTIP--VSKNDEVFNTIG 381
QY 343 SVAKTITVEYQREN--ENRTANLOSMEFTNIDTKNHT-----VEQTIYNPLRSYAKE 392
DB 382 NITTKTTANIQIPDYVVENNSIGSAFT--ETVSHVGNKKNPGYIKOTIIVNSENSLTN 439
QY 393 TNVNI-----SGNDEGSTIIDSTIIKVKYKVGNDQMLPDSNRIYDVS--EYEDVTNDYA 446
DB 440 AKLKVAOAHSSYPNNIGQINKVDTKIKYQVPGYTL--NKGVDVNTKELTDTN-QYL 495
QY 447 Q---LGNNDVNINFGNIDSPYIIKYSKDPKDKDYTTIQQVYVYMQTTINETGERT 503
DB 496 QKITTYGDNNSAVIDFGNADSAVYVMVNTKFOYTNSESPTLVOMATLSTSTN-----KSV 549
QY 504 SYDNTIAFSTSSGOGGDLPEPEYKIGDYVMEVDVDCIGQNTNNDKPSLNVLTLYTP 563
DB 550 STGNALGFTNNOSGAG-----QEVYKIGNVYMEDTKNGVDEL--GEKGVGNTVTVY-FD 602

```

```

QY 564 DGTSKSVR---TDEDGKY 578
DB 603 NNTNTRKVGCAVTRKEDGKY 620

RESULT 8
099M47
ID 099M47 PRELIMINARY; PRT: 1385 AA.
AC 099M47;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
  protein.
GN SDRD OR SAV0562 OR SA0520.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
  Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohba T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
  Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
  Kanemori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
  Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
  Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
  Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
  Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
  aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003359; BAB56724.1;
DR EMBL: AP003131; BAB41751.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS: TIGR01167; LPTXG_anchor; 1.
DR TIGRFAMS: TIGR01168; YSIRK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 1385 AA; 149643 MW; 2A0C2277733B3C1D CRC64;

Query Match 18.6%; Score 559; DB 16; Length 1385;
Best Local Similarity 28.8%; Pred. No. 1,3e-13;
Matches 177; Conservative 117; Mismatches 235; Indels 86; Gaps 27;

QY 1 SDEEKNV---INNOSTIDDDNNQIIKKEETNNNDGIEKREDETESTNVDEN-NEAT 56
DB 55 STNKELENTTASDNOSSDKVDMQOLNODNTRKNOKEMVSSQGNETTSGNKLIEE 114
QY 57 FLOKTPQDNTLHEEVEKSSSVSSSSSIDTQAQPSHTTINREESVQTSNVEDSHVD 116
DB 115 SVOSTGNKVEVSTAKSDQASPKSTNEDLNTRQ-----TISNOELQ-PDLGKNSYVN 168
QY 117 FANSKIKESNTESGKEENTIEOPNKKEDSTTSQPSGYNTI-DEKISNOELL-NLPIN 173
DB 169 -----VQPTNEBKRVKVDAKTESTTLNVKSDAISK-----NAETLVDNNSNSNN----- 211
QY 176 ENKARPLSTTSAPSIKRVVNOLAAEOGS-----NVNHLKLYDQSTBEYDSEGVIRA 231
DB 212 ENNADITLPKSTAPKSLNTRMAAALQPNSTDSKKNVNDLITSNTTLTYVVDADKNNKIY 271
QY 232 HDAENLIYDVFEEVDKVKSGDTMTVDIDKNTV-----PSDLTDSFTIPRIKD-NSG 284
DB 272 QDYLSLKSQIT--VDKVKSGDYFTIKY-SDTYQVYGLNPEIDIKN--IGDIKPNNG 325
QY 285 IATGTYNKKKQITFTFDYDKYENIKAHKLTSYIDSKSVPPNNKTLDEYKTAISV 344
DB 326 IATAKHDTANLLITFTFDYVDRENSVOMGINSIYMDADTIPVD--KKDVPSPSVIGNO 383
QY 345 NKITTYEYQRP--NENRTANLOSMEFTNIDTKNHT-----VEQTIYNPLRSYAKE 394

```

```

Db 384 ITTTADITPPAYKADNNISGSAFT--ETVSHVGNVEDPGYVNVVYVNPMDKDLGK 441
QY 395 VNISGNDEGST----IIDSTIIIVYKVGDNQNL.PDSNRIDY--SEXEDVTD--DYA 446
Db 442 LKVEAYHPKPYPTNIGQINQNMVTKIRYVEGYTL---NKGYDVMTNDLVDVDEFFKNM 498
QY 447 QUGNNNDVINGNIDSPILIKIVISKYDKNDDYTTIOQTVMQTTINEYGEFTASD 506
Db 499 TYGNSQVNLDFGDTISAVVNVNTRFOYTNESEPTLVQMATLSSTGN-----KSVSTG 552
QY 507 NTIASTSSGOGGDLPEPKTYKIGDYVEDVDKDIQNTNNEKPLSNVLTLYTPDGT 566
Db 553 NALGFTNNQSGAG---QEVYKIGNYWMEDINKNGVQEL--GKGVGVVTVV-FDNNT 605
QY 567 SKSVR---TDDEGKY 578
Db 606 NTKVGEAVTKEDGSY 620

```

## RESULT 9

```

ID 09K114 PRELIMINARY: PRT: 1733 AA.
AC 09K114;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative cell-surface adhesin SdrF.
GN SDRF.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9491;
RX MEDLINE=20340957; PubMed=10878118;
RA McCrea K.W., Hartford O., Davis S., Ni Eidhin D., Lina G.,
RA Speziale P., Foster T.U., Hook M.,
RT "The serine-aspartate repeat (Sdr) protein family in Staphylococcus
RT epidermidis."
RL Microbiology 146:1535-1546(2000).
DR EMBL: AF245041; AAF72509.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
Q SEQUENCE 1733 AA; 184720 MW; D8D62EA1692FD4E3 CRC64;

```

Query Match 16.6%; Score 499; DB 2; Length 1733;  
 Best Local Similarity 27.1%; Pred. No. 2.8e-11;  
 Matches 168; Conservative 104; Mismatches 275; Indels 72; Gaps 25;

```

QY 1 SSDEKNDVYNNQSDINTDNNQIKKEETNNYDGIKREKSEDRCTESTVNDENEA-TFLQ 59
Db 148 TSTTQODSTREKNNPSLKLNDLSSSTSKTDEHSTKQAQST--NKSVLDTNDSPQSE 206
QY 60 KT-PQDNTHLTTEEVKKESSVSSMS-----SIDPAQPSHTTTN-REESVQTSN 108
Db 207 KTSQDANNSTQNSQAPSKQSDSKSEOKVYTKFNDDEPTQDVEHTTTKLKPPSVSTDS 266
QY 109 VDSHVSDFPANSKIKESNTESGKEENTIQPKVKEDSTTSQSGTNDIKI----- 161
Db 267 VNDK---QDYTRSAV-----ASLGVDSNETEATINAVRDNLDKAFSREQINEAITALEALK 320
QY 162 --SNODELNLDP--INEXENKARPLSTSAQPSIKRVYNQALAE--QGSNNHLIKYVD 215
Db 321 DSNPDYGVDPDLALNRSQNSP--HKSASP---RMNLSLAEPNSGKNVNDKVKITN 375
QY 216 Q--SITEGVDSEGYIKAHDAENLYDVFVEVDKVKSDTNTVDIDKNTVPSDLTDSFT 273
Db 376 PTLSTNKSNNHANNVIMPTSNQFNLKANYELDSDIKEGDPTFIKYGQYIRPGGELPAI 435
QY 274 IPIKDNSEIATGTGYDKKKNQITTFPDYDKYENIKAHKLKLSYIDKSAVPPNNNTKL 333

```

```

Db 436 KTLQASKDSIYANGVYDKTTNTTTFYTFNYDYQNTIGSFDLITPKRETAINDQNY 495
QY 334 DVEYKTAASSVNRKTTVEYQRPENENRTANLQSMFTNIDTKHTVBOYIINPLRSAKET 393
Db 496 PKEVTIANEVYKKDFIVDGNKKDNTT--AAVANVDVNNKHHDEVYVLLNQNQNPXA 552
QY 394 NVNISGNDEGSTIIIDSTI--IKVYKVDNQLPDS-NRIYDSEEDVTDNDYQAL- 448
Db 553 KY-----FSTVANGEIPGEVRYEYTDINAVDSNPDLNSNNKDYTSQAPKVS 604
QY 449 GNNNDVNIENFNGI---DSPEYIKIVISKYDKNDDYTTIOQTVMQTTINEYGEFTASY 505
Db 605 ADGTAVDINFANSMANGKKYITQAVRPGTGNVTT--EYMLTRPGTTN--TNDYRGK 660
QY 506 DNTIASTSSGOGGDLPEPKTYKIGDYVEDVDKDIQNTNNEKPLSNVLTLYTPDGT 565
Db 661 STTVYLYNGSSYTAQGDNP---TSLGDYVWLDKNNQVQ--DDDEKGLAGVYTL--KDS 713
QY 566 TSKS---VRTDEGKYQED 581
Db 714 NNRELQRYVTTDOSGHYQED 732

```

## RESULT 10

```

ID 099W48 PRELIMINARY: PRT: 953 AA.
AC 099W48;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Ser-Asp rich fibrinogen-binding, bone sialoprotein-binding
DE protein.
GN SDR OR SAV0561 OR SA0519.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iano J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003359; BAB56723.1;
DR EMBL: AP003131; BAB41750.1;
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LpxTG_anchor; 1.
DR TIGRFAMs: TIGR01168; YsirK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
Q SEQUENCE 953 AA; 103292 MW; 729A169A074A1A5 CRC64;

```

Query Match 16.2%; Score 489; DB 16; Length 953;  
 Best Local Similarity 28.2%; Pred. No. 3.7e-11;  
 Matches 162; Conservative 82; Mismatches 236; Indels 94; Gaps 22;

```

QY 35 GIEKSEDRTESTVNDENEAFLQKTPQDNTHLTEEVKKESSVSSSSSIDTQAQPSH 94
Db 45 GHEAKAAETHTNELQSKNETT---APSENK--TTEKV--DSRLDKDNTQTATADQPPV 96
QY 95 TTNREESVQTSQDANVEDSHVSDFPANSKIKESNTESGKEENTIEOPKNKVEDSTQSPSGY 154
Db 97 T-----WSDSATVKEETSSNMQS-----PONATASQSTQTSNV 129

```

```

QY 155 TNIDEK---ISNDELINLPINEYEN-KARPLSTSAOPSIKRYTVNOQA-EOGSNVNH 209
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 TTNKDSSTYSNETDKSNL---TQAKNVSTPKTTTITKORALNRAVNTVAAPQOGIVND 187
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 210 LIKTPDSTI-----TEGYDSEGVAKHADEANLTYDVFEVDKRVSGDTMT 256
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 KVFETNIDIALDKGVAKTKGTGNFPMATSSDVLK-----LKANTTIDSVKEDGTF 239
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 257 VDIDKNTVPSDLADSTFTPIKIDNGSEIATGYDNKNKQITVTFYVDKYENIKAKHLK 316
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 240 FRYGQYRPGSVRLPQGTQNLNNOGNILAKGIYDSKNTTTFYFVVDQTVNVSSE 299
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 LTSTIDSKYPPNNNTKLDVEKTAALSSVNTTYEYQRPNNKRAANQSMETNIDTKNHT 376
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 QVAFKAKRENATTDKTAKEVETLGNPTYSKDVIYD---GNCKGQOOLISSTNYINNEDEL 356
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 377 VEQITVYN-PLRYSAKETNV-NISGNDEGSTIIDSTIIKRVKVDGNQNLPSDNRIYD 434
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 RNMTVYVNOQPKTYTKETFTVNTL---GKFNPDAKNKETIYETDQNOQVDS-FTPDY 410
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 435 SEVEDVTND-DYAQLGNNDVNIHFGN---IDSPYIIKVISKYDPNKKEDYTTIQCFVTM 489
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 SKLKDVTGQPDVYISNDNKRATFDVLLNGOSSDKOYIIQVAYPDSSTGDKIDYLET 470
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 QTTINEYGEFRTASYNNTJAFSTSSGOGGDLPEPKTYKIGDIYVEDVDKDGIONTNDN 549
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 QNGKSSMSN-----SYSVNGSSTANGD-----QKKRYNGDYVWEDTNRKDGODA--N 516
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 550 EKPLSNVLYTLTPDGTG-KSVRTDEGKYOFDG 582
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 517 EKGIGVYVLKDSNGKELDRFTTDENGKIQFTG 550
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q93MH7 PRELIMINARY: PRT: 881 AA.
ID Q93MH7
AC Q93MH7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Receptin Fbl precursor.
GN FBL.
OS Staphylococcus lugdunensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=28035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2342;
RA "A. fliorinogen-binding protein of Staphylococcus lugdunensis.";
L Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF040483; AAK85649.1; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LpYtg_anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 39 POTENTIAL.
FT CHAIN 40 881 RECEPTIN FBL.
SQ SEQUENCE 881 AA: 94251 MW: 24296C4959C4F19B CRC64;

Query Match 15.3%; Score 460.5; DB 2; Length 881;
Best Local Similarity 27.3%; Pred. No. 3.9e-10;
Matches 154; Conservative 88; Mismatches 231; Indels 91; Gaps 23;

QY 6 KNDVNNOSINTDNNQIKKEETNNVIGIEKRSEDRTESTTVDNEATFLQK--TPQ 63
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 KQOIGNNDA--TGDP-----QODNNYN-----NEISNOEATQONKOITOS 86
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 DNTHTLEEVKSSSVSSNSSIDTAQOPSHITTINRESVQTSIDNVEDSHVDFANSKIK 123

```

```

Db 87 DNVNSEAOAINELSD-----SHRTVNK--ATEALDNNSTLMTSTVSPATK 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 ESNTESK--EEN---TTEOPNKYKEDSTISQPSGYNINDEKISNO--DELLNLPINEYE 176
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 QDPTTSQOTQOENNDAATQRTKNTYKQDGNNNVLSQVATFNDQSSNOPRSHLNTSTVYN 190
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 N--KARPLSTSAOPSIKRYTVNOQAEOGSNVNHLIKVTIDQSTIEGYDSEGVAKHA 234
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 NNHOVRLAKVEATNTNNNTQTSIDSNKLSNVTATEAD-----TTPHKA 238
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 ENLIYDTFEYVDKRVSGDPTWTVIDRNTVPSDLSTFPIKIDNGSEIATGYDNKN 294
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 EYVLANRFRQAPDPVQAGDSIKITIPQALNNGVATAKAPNIWAG-QIATGTIDEG 297
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 295 KQITFTFDYVDKYENIKAKHLKLTSYIDSKSVPPNNNTKLDVEYTAALSSVNTTYEYOR 354
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 NLI-YTFETDVTNNKNTIGISIPGYIDPKVHTGT-KVNLSTIGQTTAKKATVYDEK 355
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 355 PNEKRTANLOSMFTNIDTKNHTVEQTIYNPLRYSAKETNVN--ISGN--GDGSGFTID 409
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 YGEFRNLSIKGTIDQIDKVNNTYQTVYVNP-----SDTVVDPYLRGSGIPGTNSNVIID 411
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 D-STIIKVKYKVDGNQNLPSDNRIYDSEVEDVTNDYAQLGNNDVNIHFGN---NIDSP 464
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 EONTSIKYKYEKAHLTDSYV-DPSNYEDVTSVKITPEKGIYQINFTEDDQINSP 470
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 YIKVISKYDPNKKDYTTIQQVTMTTINEYGEFRTAS--YNTJAFSTSSGOGGDL 522
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 YVVIINGHVDPNSNG-----NLVLRSLYLYDSNFTVSNAMNEVEYHAGSNGCIGID 524
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 523 PPEKTYKIGD-----YVWEDVDKD 541
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 KVIIPDQDSDSDAYSDDADSD 548
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q86487 PRELIMINARY: PRT: 947 AA.
ID Q86487
AC Q86487;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Src protein.
GN SDR.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEWMAN;
RA MEDLINE=99098700; PubMed=9884231;
RA Josephson E., McCrea K., Ni Eidhin D., O'Connell D., Cox J., Hook M.,
  Foster T.J.;
L "Three new members of the serine-aspartate repeat protein multigene
  family of Staphylococcus aureus.";
DR Microbiology 144:3387-3395(1998).
DR EMBL: AJ005645; CAA06650.1; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR TIGRFAMs: TIGR01167; LpYtg_anchor; 1.
DR TIGRFAMs: TIGR01168; YSIRK_signal; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 947 AA: 10288 MW: 306EPD6E35121554 CRC64;

Query Match 15.1%; Score 455; DB 2; Length 947;
Best Local Similarity 27.2%; Pred. No. 6.8e-10;
Matches 158; Conservative 93; Mismatches 222; Indels 108; Gaps 25;

QY 35 GIEKSEDRTESTTVDNEATFLQKTPDQNTHTLEEVKSSSVES-----SNSIDTAQ 90
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 GHEAKAEHTNGLNDSKNHET-----APSEN-----KTKRKVDSFQKDNQTAQATAD 92

```

```

QY 91 QPSHTTINRESVOTSDVNEHSHVDFPANSKIKESNTESKEENTIEOPNKVKEDSTISQ 150
DB 93 QPKVT-----MSDQATVETSSNMOS-----PONATANOSTTK 125
QY 151 PSCVTINDEK---ISNODELLNLPINEYENKARPLSTSAOPSIK-----VTVNOLAA- 201
DB 126 TSNVTINDKSTTYSNEDKSNL-----IOAKDVSTTPKTTIKPRLNMAVNTVAAP 179
QY 202 EOGSNVHLIKVT--DOSTIEGY--DSEGVIK--AHDAENILYDVEFVDKVSQDGMT 256
DB 180 QOQTNVDKHFNSNIDIAIDKGVNOTGTKEFWATSDVLKANYTIDDSVKEGDTFT 239
QY 257 VDIKNYPSDLTDSFTIPKIKDNGSELIATGYDNKKOQITTEPTDQVDEYENIKALK 316
DB 240 FKYOQFRRPGSVRLPSOTQNLVNOGNTIAGITDSTNTTTFYTNVDTYVRSFE 299
QY 317 LTVYIDSKSVNNNTKLDVEYKTAISSVNTIYEOHPNENRANTANOSMTNIDTKNHT 376
DB 300 QVAKAKRNATTDKTAVKMEVTLGNDYSEELIYD---GNKKAQPLISSNTYINNEELS 356
QY 377 VEOITLYN-PLRYSAKETNV-NISGNGDEGTIIDSTIIVYKVGONQNL-----PDSN 429
DB 357 RNMTAYVNOQPKNTYTKQTFVTNLT-----GYKFNPNNAKFKIYEYTDQNOCFVDSFTPTS 411
QY 430 RIYDYSEEDVTNDYDQAGLNNDVNFNGIDS--PYIIKVI-----SKYDPNKKDQYTT 482
DB 412 KLDVTDQFDVI---YSDNKTATVLDLKKGOTSSKQIIOQVAVPDNSSTDNQIDY-- 466
QY 483 IOQVTVMQTTINEYTGERTASYDNTIAFSTSGOGQDLPPEKTYKIGDYWEVDKDG 542
DB 467 -----TLDTDKTKYSW---SNSYSVNGSSSTAND-----QKKYNLGDYVWEDTNNKG 511
QY 543 IQNTNNEKPLSNLVLTITYPDGTG--KSVRTDEGKIOPDG 562
DB 512 KQDA--NEKGIGYVYLKDSNGKELDRTTDENCKYQFTG 550

RESULT 13
QY 053653 PRELIMINARY; PRT; 933 AA.
AC 053653;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Clumping factor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEWMAN;
RC MEDLINE=94224142; PubMed=8170386;
RA McDevitt D., Francois P., Vaudaux P., Foster T.J.;
RT "Molecular characterization of the clumping factor (fibrinogen receptor
RT of Staphylococcus aureus."
RL MOL. MICROBIOL. 11:237-248(1994).
EMBL Z18852; CAAT9304.1;
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMs: TIGR01167; LPXTG_anchor.1.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 933 AA; 97058 MW; EB51A6DEZFF759F4 CRC64;

Query March 15.0%; Score 452; DB 2; Length 933;
Best Local Similarity 26.6%; Pred.No. 8.7e-10;
Matches 154; Conservative 105; Mismatches 234; Indels 86; Gaps 26;

QY 1 SDEKENDVYNNOSINTDNNQIKKETNNYDGIKESRSDRESTTNDVENATFLQK 60
DB 51 SNEKSNDSSSVSAPKRTDNTN-----SDTKSSNTN--NGETSVAQ 91

```

```

QY 61 TPQONHLLTEEVKESSESVSSNSIDTAQOPSHTTINRESVOTSDVNEHSHVDFPANS 120
DB 92 NP-----AQOETOSSSTNATTEETPVGTGATTTTNOANTPATPOS--SNTAMELVN- 143
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTSOPSGVTINDEKISNODELLN-PIPIYENKA 179
DB 144 --QTSNETTFNDTIV-----SSVNSPONSTNAENVSTTQDTSLEATPSN---NES 189
QY 180 RPLSTTSAPSIKRYTVN-----OLAAEOGSNVHLIKVTDQ--SITEGYDSEGV 228
DB 190 APQSTDSNKNQVNOAVNTSAPRRRAFSLAADAAPAGATDITNQLTNVTGVI--DSGTT 248
QY 229 IKADAENILYDVEFVDKVSQDGMTVDDKNTVPESDLTDSFTIPKIKDNGSELIATG 288
DB 249 VYRQAGVYKLNKNGFSPNSAVKQDTEKITYPKELNLNGVSTAKVPPIMAGD--QVLANG 307
QY 289 TYDNKKQITTYTIDYDYKYNIRAKHLKLTYSIDSKVPNNNTKL-DVEYKTAISS--VN 345
DB 308 VIDS--DGNVIYITFDYVNTKQDVAKATLTPAYID-----PENVKKGNVTLATGIGSTTAN 362
QY 346 KTIYEOHPNENRANTANOSMTNIDTKNHTVEQTIYNPL--RYSAKETNVNISGNGDE 403
DB 363 KTVLADVEKYGKFFNLSTIKGTIIDQIDKTNNTYRQTIYVNPBGDNVIAVLGNLKPNTDS 422
QY 404 GSTIIDSTIIVYKVGONQNLPSNRIYDYSEEDVTNDYDQAGLNNDVNFNGIDS----- 459
DB 423 NALIDQOQTSIKYKVDNAADLSESYFV--NPENEDVYNSVNTFFPNQKVEENTDDP 481
QY 460 NIDSPYIKYVSKDPN-KDOYTTIQOVTVMQTTINEYTG--FRTASYDNTIAFSTSG 516
DB 482 QITPIYIVVNGHIDPNKGD-----LALRSTLYGNSNINIMSMWDEVAENNSG 534
QY 517 QGQG-DLP--PEKTYKIGDY--VWEDVDK-D-GIQNTNDN 549
DB 535 SGDGIDKRVVPEQDEPEIEPIPEDSDSDPGSDGSDS 573

RESULT 14
QY 0932C5 PRELIMINARY; PRT; 935 AA.
AC 0932C5;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Fibrinogen-binding protein.
GN FNB OR SAV0811.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Ihan J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
EMBL AP003360; BAB56973.1;
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR001899; Gram_pos_anchor.
DR TIGRFAMs: TIGR01167; LPXTG_anchor.1.
DR TIGRFAMs: TIGR01168; YSIRK_signal.1.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 935 AA; 96950 MW; DC5A2D92CE3BA91C CRC64;

```

Query Match 15.0% Score 452; DB 16; Length 935;  
Best Local Similarity 25.9%; Pred. No. 8.7e-10;  
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```
QY 1 SSDEKNDVINNNOSINTDDNNQIIKKEETNNYDIEKRSRDRRESTINVDENATFLQK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 51 SNEKSDNDSSVSAPKTDITN-----SDTKTSSNTN---NGETSVAQ 91

QY 61 TPQDNTLHEEVEKSSSVSSNSIDTAQOPSHTTINRESVOTSDNVEDSHVDFANS 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 NP-----AQOETQSSSTNATTEETPYTGATTTTQANTPATQOS-SNTNAEELVN- 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTTSQPSGYTINDEKISNODELIN-LPINEYENKA 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 144 -QTSNETTSDNTMTV-----SSVNSPQNSTMAENVSTQDTSTEARPPSN---NES 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 RPLSTTSA-----OPSIKRVTVNOLOA-----EQGSNVNHLIKVTDQSITEGYDD 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 APQNTDASNKDVSOAVNPSTPRKMAFSLAADAAPAGTDTTN--QITDVKVT---ID 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 SEGVIKADHAENLIYDVFEEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEI 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 SGTIVYPHQAQYVNLNGFSPNSAVKGDFTKIVPKELNLNGYSTAKVPPIMAGD-QV 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 IATGYDNKKNQITFTFDYVDKYEENIKAKLKLTSYIDSKVPPNNNTKL-DVEYKTAAL- 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 304 LANGVIDS-DGNVIYTFDYNKKNENYANTIMPAYID---PEWTKTGAVTLTGTIGT 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 342 SSVKKTIVYQRPENENTANLOSMTNIDTKNHTVEGTIYNPLRSKAKETNV----- 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 359 NTASKTVLIDYEKGQFNLISIKGTIDQIDKTNNTYROTIIYVNP-----SGDNVLPALY 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 -NISNGDEGSTIIDDSTIIKVKYKQGNQNL.PDSNRIDYSEYEDVNDQVLAQGNNDV 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 414 GNLPNTKSNALIDAKNTDIVRYR-DNANDLSESYVNPDSFEDVTQVRISEFPNANQY 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 NINF-----GNIDSPYIIKIVISKYDPNKDDYTTIQOTVMTQTTINEYTGEEF--RTASYDNT 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 473 KVEPPTDDQITTPYIVVNGHIDPASTG-----DLALRSTFYCYDSENFIRMSKMWNE 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 509 IAFSTSSGQGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 549
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 VAFNNGSGSGDIDKPVVPEQDEGEIEPIPEDSDSPGSDSGSDS 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 15

Q99VJ4 PRELIMINARY; PRT; 989 AA.

```
AC Q99VJ4;
T 01-JUN-2001 (Tremblrel. 17, Created)
01-JUN-2001 (Tremblrel. 17, Last sequence update)
01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fltrogen-binding protein A, clumping factor.
GN CLFA OR SA0742.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsunaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-U Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hiraoka H., Kihara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003131; BAB41975.1; -
```

DR InterPro: IPR000515; BPD\_transp.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR Trifam: TIGR01167; LpxTG\_anchor. 1.  
DR PROSITE: PS00402; BPD\_TRANS\_PEN\_MBR; UNKNOWN\_1.  
DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 989 AA; 102407 MW; DA6F807539623467 CRC64;

Query Match 15.0% Score 452; DB 16; Length 989;  
Best Local Similarity 25.9%; Pred. No. 9.2e-10;  
Matches 152; Conservative 109; Mismatches 224; Indels 102; Gaps 27;

```
QY 1 SSDEKNDVINNNOSINTDDNNQIIKKEETNNYDIEKRSRDRRESTINVDENATFLQK 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 51 SNEKSDNDSSVSAPKTDITN-----SDTKTSSNTN---NGETSVAQ 91

QY 61 TPQDNTLHEEVEKSSSVSSNSIDTAQOPSHTTINRESVOTSDNVEDSHVDFANS 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 NP-----AQOETQSSSTNATTEETPYTGATTTTQANTPATQOS-SNTNAEELVN- 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KIKESNTESGKEENTIEOPNKVKEDSTTSQPSGYTINDEKISNODELIN-LPINEYENKA 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 144 -QTSNETTSDNTMTV-----SSVNSPQNSTMAENVSTQDTSTEARPPSN---NES 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 180 RPLSTTSA-----OPSIKRVTVNOLOA-----EQGSNVNHLIKVTDQSITEGYDD 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 190 APQNTDASNKDVSOAVNPSTPRKMAFSLAADAAPAGTDTTN--QITDVKVT---ID 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 SEGVIKADHAENLIYDVFEEVDKVKSGDTMTVDIDKNTVPSDLTDSFTPIKIDNSGEI 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 SGTIVYPHQAQYVNLNGFSPNSAVKGDFTKIVPKELNLNGYSTAKVPPIMAGD-QV 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 IATGYDNKKNQITFTFDYVDKYEENIKAKLKLTSYIDSKVPPNNNTKL-DVEYKTAAL- 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 304 LANGVIDS-DGNVIYTFDYNKKNENYANTIMPAYID---PEWTKTGAVTLTGTIGT 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 342 SSVKKTIVYQRPENENTANLOSMTNIDTKNHTVEGTIYNPLRSKAKETNV----- 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 359 NTASKTVLIDYEKGQFNLISIKGTIDQIDKTNNTYROTIIYVNP-----SGDNVLPALY 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 -NISNGDEGSTIIDDSTIIKVKYKQGNQNL.PDSNRIDYSEYEDVNDQVLAQGNNDV 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 414 GNLPNTKSNALIDAKNTDIVRYR-DNANDLSESYVNPDSFEDVTQVRISEFPNANQY 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 NINF-----GNIDSPYIIKIVISKYDPNKDDYTTIQOTVMTQTTINEYTGEEF--RTASYDNT 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 473 KVEPPTDDQITTPYIVVNGHIDPASTG-----DLALRSTFYCYDSENFIRMSKMWNE 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 509 IAFSTSSGQGG-DLP--PEKTYKIGDY--VWEDVDKD-GIONTNDN 549
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 527 VAFNNGSGSGDIDKPVVPEQDEGEIEPIPEDSDSPGSDSGSDS 573
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: March 14, 2003, 13:00:05  
Job time : 27.2509 secs